



Dipl.-Ing. Stefan Lengauer, BSc BSc

Crossmodal Search, Visual Exploration, and Restoration of Digital Pottery Artifacts

DISSERTATION

to achieve the university degree of
Doctor of Technical Sciences

submitted to

Graz University of Technology

Supervisor

Univ.-Prof. Dr. Tobias Schreck

Institute of Computer Graphics and Knowledge Visualisation

Co-Reviewer and Examiner

Jun.-Prof. Dr. Hubert Mara

Institute for Computer Science

Martin-Luther-University Halle-Wittenberg Graz, November 2022

ABSTRACT

One of the most important archaeological cues for the study of ancient civilizations is pottery, which makes up by far the largest group of excavated artifacts, since every new excavation effort potentially yields up thousands of fragments. Similar to the humanities in general, this research area, belonging to the field of Cultural Heritage (CH), experienced several ambitious digitization efforts in the last two decades. The resulting digital surrogates enable the application of computerized tools, which can support tedious manual processes. 3D, in particular, has become a state-of-the-art data modality for the documentation of numerous types of exhibits. This can be attributed to both the more easily affordable capturing hardware and freely available software suites for the easy archiving and analysis of 3D data. However, not only newly unearthed objects but also decades old documentations of excavated ceramics have been digitized to a large extent and made publicly accessible via online repositories and collections.

On the downside, the efficient browsing of such repositories as well as – one of the most important use cases for archaeological research – the content-based searching for object similarity is seldomly supported. Instead, most platforms allow to filter their collections based solely on associated metadata. As this information is incomplete, erroneous, or missing altogether for many entries, a potentially large number of relevant records could easily be overlooked. Also, this type of querying could result in an unmanageably large result set (without any relevance ordering), if those properties cannot be determined selectively enough. This stresses the need for specialized retrieval systems, which allow posing queries in a content-based manner, meaning that the data itself, *e.g.*, an image or 3D model, is the basis for a search. This type of retrieval has been researched for decades, scenarios where the query and the search space exhibit different data modalities, *e.g.*, 3D model vs. profile curve, is still a sparsely addressed problem. As such scenarios are, however, particularly relevant for researchers of the CH domain, respective retrieval systems have to be capable of both supporting different modalities and bridging modal boundaries.

Orthogonal to the problem of content-based searching is the visualization and interactive exploration of (large) object collections with the goal of unveiling similarity clusters. Respective techniques allow an archaeologist to detect cross-correlations among object traits; discover (yet unknown) links, since related objects, or even parts of the same object, are often scattered across different museums; and generate hypotheses regarding ancient trade routes or workshops. At the same time, visualization enables the exposition and presentation of CH artifacts as virtual museums and exhibitions.

Finally, digital reconstruction approaches support the efficient generation of shape and texture hypotheses by automatizing the workflow of an archaeologist who leverages both a vessel's symmetry properties and sequential structures to approximate missing geometry and surface painting. Digital methods are also a viable alternative to physical reconstruction, which is most of the time not feasible due to cost and labor considerations or not even allowed because of the fragile nature of CH artifacts.

This thesis provides contributions to all the three mentioned archaeological processes with computer-aided techniques, relying on established as well as newly developed techniques of computer vision and knowledge visualization. Within the context of content-based retrieval, a new query modality (3D+Sketch), particularly suited for the search based on incomplete pottery artifacts, is presented. Newly designed annotation workflows, benchmark datasets, and evaluation metrics pave the way for both the training of learning-based retrieval techniques and the comparative evaluation of methods. In terms of visualization, a novel interactive linked view exploration system supports the uninterrupted exploration of annotated datasets concurrently along different object traits. Finally, in terms of reconstruction, the textual completion of worn off surface paintings is studied. To this end, a novel approach is presented which, first of all, determines the generation grammar of an ornament sequence in order to detect gaps and fill them with plausible texture. All of these techniques are able to reduce the required labor *w.r.t.* their currently used purely manual counterparts significantly. Their viability and relevance for archaeological use cases has been continuously evaluated with collaborating archaeologists who also aided in the design and execution of both qualitative and quantitative evaluations. Even though the proposed concepts are experimental, they will hopefully find their way into the everyday archaeological use.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, Tobias Schreck, for his continuous support and guidance throughout the years I was able to spend at the Institute of Computer Graphics and Knowledge Visualization (CGV). His door was always open for feedback and fruitful discussions.

I also want to thank my other colleagues at the CGV with whom I worked with on several projects and published numerous papers. In particular, I want to mention Reinhold Preiner, who acted as a mentor and was an indispensable support in the conceptualization and writing of my first research papers.

In particular, I want to mention Reinhold Preiner, who acted as a mentor and was a indispensable support in the conceptualization and writing of my first research papers. I am also grateful that I had the honor of working with Alexander Komar and Peter Houska on several occasions, which always resulted in valuable results.

Furthermore, I would like to thank all the external collaborators who contributed with significantly to my scientific work: Stephan Karl and Elisabeth Trinkl from the University of Graz, as well as Ivan Sipiran and Benjamin Bustos from the University of Chile.

Finally, I want to thank my friends and family who supported my endeavours in a different manner. My warmest thanks to my parents Elisabeth and Gerhard Lengauer who inspired me to work hard, as well as to my beautiful girlfriend Marlene Findenig, who never gave up on me and managed to endure me for such a surprisingly long time.

Once more thanks a lot to all of you!

AUTHOR'S DECLARATION

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

Date

Signature

LIST OF PUBLICATIONS

WORKSHOP PUBLICATIONS

- **Lengauer, S.**, Komar, A., Labrada, A., Karl, S., Trinkl, E., Preiner, R., Bustos, B., Schreck, T., “Sketch-Aided Retrieval of Incomplete 3D Cultural Heritage Objects”, in *Eurographics Workshop on 3D Object Retrieval*, The Eurographics Association, 2019, pp. 17–24. DOI: [10.2312/3dor.20191057](https://doi.org/10.2312/3dor.20191057)
- **Lengauer, S.**, Komar, A., Labrada, A., Karl, S., Trinkl, E., Preiner, R., Bustos, B., Schreck, T., “Motif-driven retrieval of greek painted pottery”, in *Eurographics Workshop on Graphics and Cultural Heritage*, Eurographics-European Association for Computer Graphics, The Eurographics Association, 2019, pp. 91–98. DOI: [10.2312/gch.20191354](https://doi.org/10.2312/gch.20191354)
- **Lengauer, S.**, Komar, A., Karl, S., Trinkl, E., Preiner, R., Schreck, T., “Visual Exploration of Cultural Heritage Collections with Linked Spatiotemporal, Shape and Metadata Views”, in *Vision, Modeling, and Visualization*, The Eurographics Association, 2020. DOI: [10.2312/vmv.20201196](https://doi.org/10.2312/vmv.20201196)
- **Lengauer, S.**, Komar, A., Karl, S., Trinkl, E., Sipiran, I., Schreck, T., Preiner, R., “Semi-automated Annotation of Repetitive Ornaments on 3D Painted Pottery Surfaces”, in *Eurographics Workshop on Graphics and Cultural Heritage*, The Eurographics Association, 2020, pp. 1–4. DOI: [10.2312/gch.20201285](https://doi.org/10.2312/gch.20201285) *
- **Lengauer, S.**, Sipiran, I., Preiner, R., Schreck, T., Bustos, B., “A Benchmark Dataset for Repetitive Pattern Recognition on Textured 3D Surfaces”, *Computer Graphics Forum*, 2021. DOI: [10.1111/cgf.14352](https://doi.org/10.1111/cgf.14352)
- Houska, P., **Lengauer, S.**, Karl, S., Preiner, R., “Direct Elastic Unrollings of Painted Pottery Surfaces from Sparse Image Sets”, in *Eurographics Workshop on Graphics and Cultural Heritage*, The Eurographics Association, 2021. DOI: [10.2312/gch.20211417](https://doi.org/10.2312/gch.20211417)
- Sipiran, I., Lazo, P., Lopez, C., Jimenez, M., Bagewadi, N., Bustos, B., Dao, H., Gangisetty, S., Hanik, M., Ho-Thi, N.-P., Holenderski, M., Jarnikov, D., Labrada, A., **Lengauer, S.**, Licandro, R., Nguyen, D.-H., Nguyen-Ho, T.-L., Perez Rey, L. A., Pham, B.-D., Pham, M.-K., “Shrec 2021: Retrieval of cultural heritage objects”, *Computers & Graphics*, vol. 100, pp. 1–20, 2021. DOI: [10.1016/j.cag.2021.07.010](https://doi.org/10.1016/j.cag.2021.07.010)

*Honorable Mention in the category “Impact on Society” at the “Computer Graphics Night 2020” by Fraunhofer IGD, Darmstadt.

-
- **Lengauer, S.**, Preiner, R., Sipiran, I., Karl, S., Trinkl, E., Bustos, B., Schreck, T., “Context-based Surface Pattern Completion of Ancient Pottery”, in *Eurographics Workshop on Graphics and Cultural Heritage*, The Eurographics Association, 2022. DOI: [10.2312/gch.20221234](https://doi.org/10.2312/gch.20221234)

JOURNAL PUBLICATIONS

- **Lengauer, S.**, Komar, A., Labrada, A., Karl, S., Trinkl, E., Preiner, R., Bustos, B., Schreck, T., “A sketch-aided retrieval approach for incomplete 3d objects”, *Computers & Graphics*, vol. 87, pp. 111–122, 2020. DOI: [10.1016/j.cag.2020.02.001](https://doi.org/10.1016/j.cag.2020.02.001)
- Karl, S., Houska, P., **Lengauer, S.**, Haring, J., Trinkl, E., Preiner, R., “Advances in digital pottery analysis”, *it - Information Technology*, 2022. DOI: [doi:10.1515/itit-2022-0006](https://doi.org/10.1515/itit-2022-0006)
- **Lengauer, S.**, Houska, P., Preiner, R., Trinkl, E., Karl, S., Sipiran, I., Bustos, B., Schreck, T., “Interactive annotation of geometric ornamentation on painted pottery assisted by deep learning”, *it - Information Technology*, 2022. DOI: [doi:10.1515/itit-2022-0007](https://doi.org/10.1515/itit-2022-0007)

BOOK CHAPTERS

- Schreck, T., **Lengauer, S.**, Preiner, R., “Forschungsprojekt crossmodale suche und visuelle exploration von 3d kulturgutobjekten”, in *Forum Archaeologiae*, vol. 91, 2019
- Trinkl, E., Karl, S., **Lengauer, S.**, Preiner, R., Schreck, T., “Cross-modal search and exploration of greek painted pottery”, in *The 3 Dimensions of Digitalised Archaeology. State-of-the-art, Data Management and Current Challenges in Archaeological 3D Documentation*, forthcoming

TABLE OF CONTENTS

	Page
List of Publications	ix
1 Introduction	1
1.1 Motivation	1
1.2 Challenges and Requirements	4
1.2.1 Research Questions	7
1.3 Scope of the Thesis	9
1.3.1 The CrossSAVE-CH Project	10
1.3.2 Research Approach	10
1.3.3 Collaboration Statement	12
1.4 Structure of the thesis	14
2 Data Sources and Acquisition	17
2.1 Data Sources	18
2.1.1 Repositories	18
2.1.2 Print Volumes	19
2.2 Data Modalities	20
2.2.1 Texts	21
2.2.2 Metadata	21
2.2.3 Profile Curves/Archaeological Drawings	21
2.2.4 Photographs	23
2.2.5 3D Models	24
2.3 Acquisition	24
2.3.1 Data Extraction from Archaeological Volumes	26
2.3.2 Crawling of Web Resources	28
3 Preparation	29
3.1 Surface Unwrapping	30
3.2 Segmentation	31
3.2.1 Object Segmentation	32

3.2.2	Motif Segmentation	34
I	Shape-based Retrieval	39
4	3D-Based and 3D+Sketch-Based Retrieval	41
4.1	Introduction	42
4.1.1	Information Retrieval	43
4.1.2	Our Contribution	45
4.2	Related Work	46
4.3	3D to Normalized Profile Curve	53
4.3.1	Concept	53
4.3.2	Results	55
4.3.3	Discussion	57
4.4	Sketch-Aided 3D Retrieval	57
4.4.1	Concept	58
4.4.2	Query Preparation	58
4.4.3	Content-based Image Retrieval	62
4.4.4	Results	66
4.4.5	Discussion	77
4.5	Conclusion	79
II	Texture-based Retrieval	81
5	Motif Retrieval	83
5.1	Introduction	83
5.2	Related Work	85
5.3	Shape and Color-based Motif Search	86
5.3.1	Concept	87
5.3.2	Experimental Evaluation	89
5.3.3	Limitations and Discussion	93
5.4	Conclusion	94
6	Ornament-Retrieval	97
6.1	Introduction	98
6.2	Related Work	101
6.3	Repetitive Pattern Recognition Benchmark	104
6.3.1	Object Selection	106
6.3.2	Annotation Concept	106

6.3.3	Annotation Tools	108
6.3.4	Evaluation Metrics	110
6.3.5	Experimental Evaluation	113
6.3.6	Possible Extensions and Outlook	114
6.4	Interactive Annotation and Classification of Geometric Ornaments	116
6.4.1	Interactive Annotation	118
6.4.2	Results	124
6.4.3	Discussion, Limitations and Open Research Questions	126
6.4.4	Applications	128
6.5	Semi-automated Creation of archaeological Drawings	130
6.5.1	Concept	130
6.5.2	Self-similarity Detection	133
6.5.3	Experimental Results	135
6.5.4	Limitations and Outlook	136
6.6	Conclusion	138
III Visual Analytics		139
7	Interactive Exploration and Visualization	141
7.1	Introduction	142
7.2	Related Work	142
7.3	Linked Views Visual Exploration System	144
7.3.1	Domain Analysis Tasks	145
7.3.2	Concept	145
7.3.3	Analytical and Interactive Support	150
7.3.4	Results and Evaluation	154
7.3.5	Discussion Limitations and Open Research Questions	156
7.4	Conclusion	157
IV Digital Restoration		159
8	Colorimetric Restoration	161
8.1	Introduction	161
8.2	Related Work	163
8.3	Content-based Repetitive Pattern Completion Workflow	166
8.3.1	Generation Rule Detection	167
8.3.2	Content-based Reconstruction	169
8.3.3	Results	171

TABLE OF CONTENTS

8.3.4	Discussion, Limitations and Future Work	175
8.3.5	Applications	177
8.4	Conclusion	178
9	Conclusion and Outlook	179
9.1	Reflection on Research Questions	179
9.2	Summary of Contributions	181
9.3	Future Perspective	182
9.4	Final Thoughts	184
A	Appendix A: Mathematical Notation	185
B	Appendix B: Algorithmic Background	187
	List of Tables	191
	List of Figures	193
	List of Acronyms	197
	Bibliography	201

INTRODUCTION

Contents

1.1	Motivation	1
1.2	Challenges and Requirements	4
1.2.1	Research Questions	7
1.3	Scope of the Thesis	9
1.3.1	The CrossSAVE-CH Project	10
1.3.2	Research Approach	10
1.3.3	Collaboration Statement	12
1.4	Structure of the thesis	14

Within this doctoral thesis, I present contributions to several everyday tasks of applied archaeology. After a motivation of the relevance of digital methods for archaeology (SEC. 1.1) and a discussion of the challenges (SEC. 1.2) inherent to this field of research, I present four research questions around which the presented work is structured. A subsequent overview on the scope of the thesis (SEC. 1.3) takes a look at the general framework (scientific project, laboratory) within which the thesis was conducted, the employed methodology (research approach), and the delineation of the work contributed by various associated parties (advisor and colleagues). Whenever the words “we” and “our” are used throughout the thesis, it refers to myself as well as them. Finally, an overview of the chapters and structure of the remainder of the document is provided (SEC. 1.4).

1.1 Motivation

The potential for digital processing of Cultural Heritage (CH) data has long been recognized for a variety of applications such as anomaly detection with ground penetrating radars [1], engagement

through virtual exhibitions and museums [2–4], similarity-based retrieval and classification of objects [5], digital restoration of artifacts [6], and virtual reconstruction of ancient sites [7].

Relevance of Pottery for Archaeological Research Independent from culture or historic period, wheel-made ceramics represent by far the largest group of excavated objects, resulting from the fact that they were produced in large quantities and neither decay like textile or woodwork, nor were they recycled like metal or glass ware. The Beazley Archive Pottery Database (BAPD) [8] alone, the world’s largest archive for ancient Greek painted pottery, features a total of more than 130,000 object records, comprising 250,000 images.

The ancient Greek pottery – commonly called *vases* – is the most outstanding group of ancient ceramics and an essential research field of classical archaeology. Most vases were designed with a specific purpose in mind, such as storage, pouring, mixing or drinking, which is commonly reflected in the traditional shapes [9]. A selection of exhibits curated by the University of Graz is given in FIG. 1.1. This form of loose morphological standardization allows to classify excavated vases and analyze the continuity or change of shapes over time. Their surface decorations, so-called *vase paintings*, are not only leveraged for the dating of archaeological contexts but also shed a light on religion, economy, daily life and society of Greek and Greek-influenced civilizations in the Mediterranean. Long-lasting research has revealed different techniques and styles indicating chronological periods, workshops or painters [10].



FIGURE 1.1. A selection of exhibits from the Institute of Classics at the University of Graz. The varied shapes and relative sizes vividly illustrate the different intended purposes. (*Image source: KFU Graz*)

The Attic period from 500 to 300 BCE, for example, saw two distinct painting styles referred to as red-figured (red clay figures on black background) [11, 12] and black-figured (black figures on red clay background) [9] vases. While the prior represents the height of Greek vase painting, the latter exhibit lower quality due to them being the result of mass production [13]. Independent

from style, vase paintings from the era often illustrate mythological figures or draped youth and characteristic scenes like wedding or warrior departure. Some of them have received attention in the *Lexicon Iconographicum Mythologiae Classicae* (LIMC) [14] as well as many publications. For an overview on this subject, we refer to the works of Cook [15] or Boardman [16].

The central scene of a vase is often accompanied by various abstract patterns filling empty spaces of background or framing the motifs in the form of ornament bands or friezes around the rotation axis of pottery objects. Latter are created by the painter while rotating the pot on the potter's wheel. Popular patterns comprise atomic shapes like spirals, concentric circles, triangles, swastika, zigzags and meanders, or more complex ones like rosettes, palmettes, leaves and birds [17]. In another, more ancient artistic phase from 900 to 700 BCE vases are exclusively decorated with such patterns, coining the apt designation as *Geometric* period, because of the favored use of geometry-like ornamentations.

Digitization and Multimodality Both advancements in technology and a general push towards digitization in all fields of the humanities results in more and more CH content becoming available in a digitized manner. Not only newly excavated artifacts are captured digitally but also historical documentations, collected over several decades, are considered in large-scale digitization efforts. *W.r.t.* ancient Greek pottery, the online presence of the *Corpus Vasorum Antiquorum* (CVA) [18], a series of fascicules appearing since 1922, as well as the BAPD [8] are prominent examples. The records in such repositories comprise a variety of data modalities, such as textual descriptions, drawings (of profile curves or surface paintings), photographs, and 3D models. On the one hand, this results from the replacement of documentation techniques with more modern alternatives over the years, *e.g.*, 3D models instead of photographs. On the other hand, different modalities are unevenly well-suited for different object traits, *e.g.*, textual descriptions are appropriate for capturing an object's cultural context as opposed to 3D models. It is thus neither feasible nor beneficial to capture all the information pertaining to a certain object with a single data modality, but different modalities are optimal for different aspects.

Besides long-term preservation of structured archaeological documentation, digitization also paves the way for the application of computer-based methods for the management, sharing [19], archiving [20], analysis [21], restoration [6, 22–31], retrieval [20, 32–42], and exploration [43–51] of CH content. Unsurprisingly, such applications experienced a rapid growth with the increased availability of data through online repositories and such within the last two decades [52]. The great opportunity of digital data processing and computerized tools is that existing (mostly manual) workflows can be simplified by supporting (or altogether replacing) undemanding yet laborious repetitive tasks. *I.e.*, the strength of computers is that they are able to process large quantities of data efficiently and infallibly. Hence, all types of tasks which can be described sufficiently well through mathematical models can be supported with automation. *E.g.*, if the similarity between two objects can be formally defined, the content-based searching through (huge) data repositories containing thousands of records, can be done efficiently with specialized

retrieval systems, while the manual pair-wise comparison is not possible within a reasonable amount of time. Even though modern learning-based approaches are able to learn a human decision finding process to some extent, given that sufficient training data reflecting this knowledge is available, some processes, like interpretation and validation, do and will require human intervention. Hence, the key objective of computer-based methods is not to replace experts, but to support them in their everyday work.

1.2 Challenges and Requirements

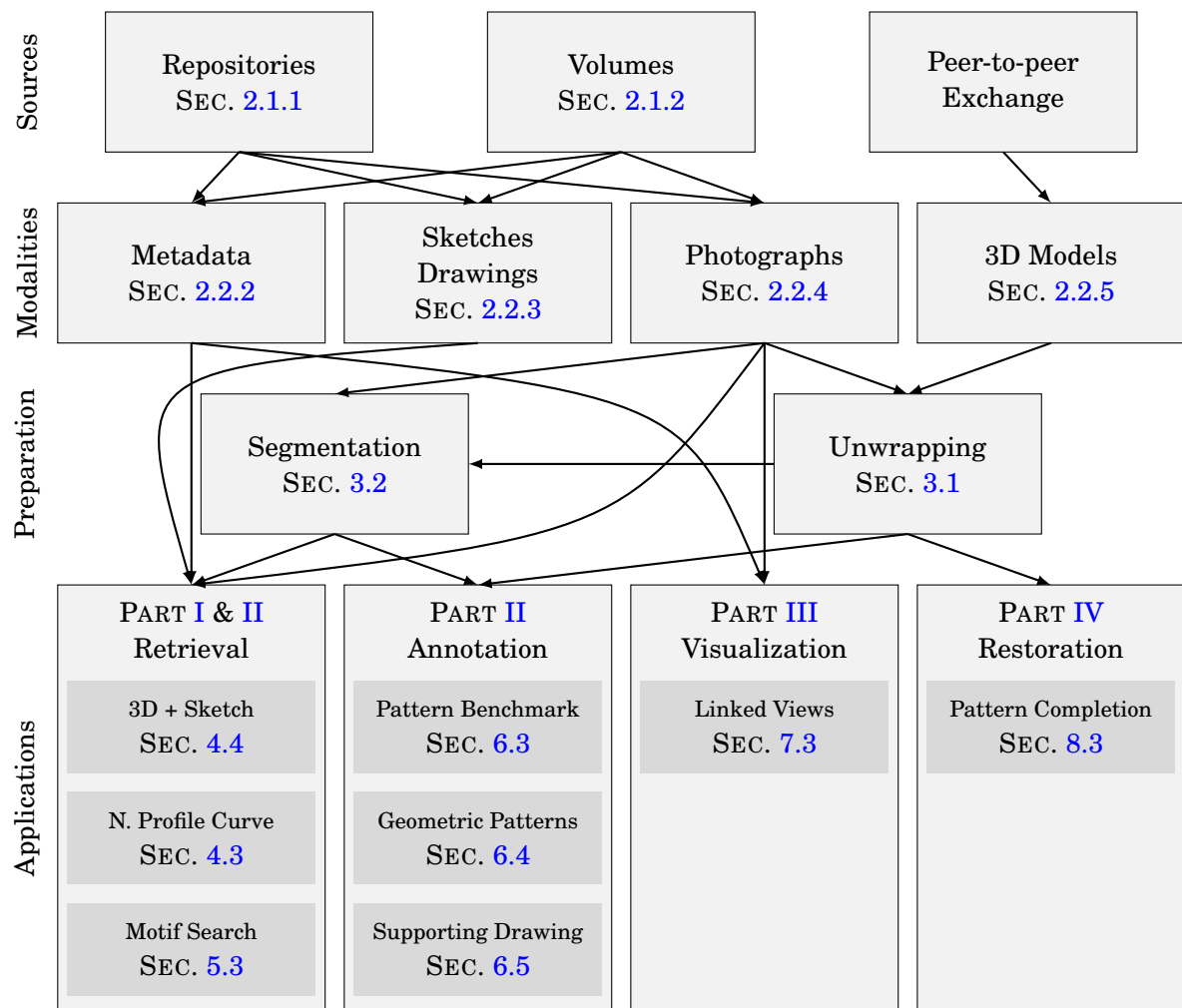


FIGURE 1.2. The experimentation with and evaluation of different application within a CH context requires relevant object data. Diverse sources can be used for acquiring said data. A particular challenge is, however, the different data modalities in which records are documented. Depending on the type of data, it has to be subjected to different pre-processing/preparation steps before it can be used for certain applications. Within this thesis, the topics of ‘Retrieval’, ‘Annotation’, ‘Visualization’ and ‘Restoration’ are touched. The references within the figure refer to the respective technical sections.

To date, most practitioners from the field of CH are concerned with the digitization of exhibits, resulting in a digital surrogate, which serves the documentation and long-term persistence of an artifact. In most cases, this remains the sole purpose, since as domain researchers are oftentimes agnostic towards – or not familiar with – the varied analysis and processing methods a digital copy enables. On the other hand – despite all the advantages of digital methods – their application for archaeological use cases is typically not straightforward. As mentioned above, a serious challenge is the variety of data modalities (see FIG. 1.2). But also heterogenous levels of quality can be observed. 3D models, for example, come in different degrees of detail, depending on both the capturing technique and the hardware they have been obtained with. They may also exhibit colorization, which can be present as texture maps, but also in the form of vertex colors in the case of high-resolution models. *W.r.t.* photographic data, researches are faced with a range of different resolutions when combining records from different repositories. Many also deviate from the proposed standards for photographing pottery objects [53, 54]. *I.e.*, some objects are photographed in front of a very neutral black or white background, while others exhibit strong gradients in the background. Some photographs are available in color, whereas others are only provided as grayscale images. Photographs scanned from fascicules often incorporate visible paper texture while digital photographs may exhibit compression artifacts. Even structured metadata is oftentimes incompatible due to the usage of different taxonomies for assigning labels, or different levels of precision for specifying dating or geographic information.

Different types of sources can be utilized to obtain relevant data (FIG. 1.2). Depending on that, researchers are faced with additional challenges. Most online resources, the state-of-the-art for data exchange, provide data in a structured and digitized manner. However, in some cases, *e.g.*, with the CVA Online [18], the data is given as a set of scanned fascicules. These scans need to be further processed (SEC. 2.3.1.1) in order to split them into textual and graphic data and link them to individual objects. The same processing steps are required for extracting object information from digitized literature, which is still, despite all digitization efforts, a non-negligible source of data. The list of data sources used within the scope of this thesis is completed by peer-to-peer exchange with fellow researchers. This type of source is especially valuable for 3D models, which are more rarely hosted online.

The ultimate goal of all acquisition efforts is to obtain relevant object data for computer-based applications. Depending on the intended application, object data needs to be subjected to different preparation steps (FIG. 1.2). For 3D models, this could comprise a so-called *unwrapping* (SEC. 3.1), a non-linear mapping of the shape’s surface into 2D space. This process allows both an uninterrupted visualization of the whole object surface and the application of image-based methods for automated analysis. Such can comprise, *e.g.*, the semantic segmentation (SEC. 3.2) of the surface in order to split an object’s colorimetric depictions into their composing parts. The same can be done for photographs, where segmentation is also used to differ between object and background pixels.

Scientific literature states different ways to group CH applications. Pintus *et al.* [21] differentiate between ‘Perception’ (enhancement of fine details), ‘Restoration’, ‘Monitoring’ (analyzing the degeneration of artwork), ‘Interpretation’ (analyzing used tools and production sequences), and ‘Collection analysis’, while Karl *et al.* [55] classify approaches *w.r.t.* the categories ‘Documentation’, ‘Retrieval’, and ‘Classification’. The structure of this thesis is based on a mixture of these taxonomies, including the broad application fields ‘Retrieval’, ‘Annotation’, ‘Visualization’ and ‘Digital Restoration’, as depicted in FIG. 1.2.

One key objective of archaeological research is the classification of artifacts. This task can be supported through similarity-based retrieval methods, which allow the targeted navigation of large collections. A non-trivial challenge is however to capture an archaeologist’s notion of similarity, which can be based on an object’s overall shape, color pattern, excavation context, *etc.*, but also a palette of miniscule details which are hard to pinpoint in a formalized manner. The fact that these notions can differ among scholars further complicates matters. *W.r.t.* pottery research, we can distinguish between two orthogonal retrieval objectives. First, the retrieval of overall similar vessels regarding shape and color patterns, *e.g.*, searching for similar black-figures lekythoi. And second, the retrieval of similar vase paintings, *e.g.*, searching for vessels depicting a similar Eros figure or wedding scene. Within the scope of this thesis, we refer to these tasks as shape-based and texture-based retrieval, respectively. However, a complete segregation of these types is not always possible since clay color and surface painting are also distinguishing characteristics for determining a shape class. Hence, some ‘shape-based’ approaches also make use of the objects’ textures. *W.r.t.* CH applications, *cross-modal* retrieval approaches – such as where the modality of a query and the search space differ – play an important role. While nowadays 3D models are the documentation method of choice, researchers also want to discover potential similarities to artifacts which have been documented by photographic or hand-drawn means which is the case of the vast majority of existing records. Closely related to retrieval is the task of annotation. With annotation, labels/classes are assigned automatically or manually to different parts of data. *E.g.*, image segmentation assigns a label to the pixel of an image. In the context of CH, researchers assign labels from taxonomies, *e.g.*, regarding the fabric, to artifacts. Retrieval methods can support such manual labeling by providing suggestions and hypotheses.

There is an increasing demand for the efficient browsing of large data collections [21]. Said can be achieved with interactive visualization and exploration techniques. The difference to retrieval is that similarities are not displayed in a ‘one-to-many’ but a ‘many-to-many’ fashion. That is, objects are not part of a linear result vector, but are visualized as data points in a high-dimensional space [50], based on different properties allowing to employ aggregation or clustering techniques. Alternatively, pair-wise similarity relations can be leveraged to arrange objects in force-driven layouts, such as graph structures [44, 45, 56] or self-organizing maps [57]. Concepts like these allow domain researchers to easily spot similarity clusters and, in turn, unveil yet unknown correlations between subsets of collections.

Lastly, digital restoration allows to restore degenerated documents or objects. One prominent field of restoration is concerned with old documents or photographs where scratches or tears are detected and replaced with meaningful texture using texture synthesis techniques [21]. In terms of objects, restoration commonly operates on 3D models. We distinguish between *reassembly*, the digital assembly of objects which are merely preserved as a set of fragments, and *completion*, the approximation of missing object parts [6]. The latter can be used to approximate both geometry [6, 22–28] and texture [29–31, 58].

1.2.1 Research Questions

Based on the aforementioned challenges of digital archaeology and under consideration of previous works in this domain, we state the following research questions, which are the core objectives we try to answer with the technical contributions presented in this thesis.

RQ1: *“How can user expertise be incorporated interactively in a query formulation for the specification of missing shape information?”*

Many retrieval approaches have been presented for the content-based retrieval of whole pottery artifacts. Most of them leverage the rotational symmetry of such objects and base their similarity computation on the comparison of profile curves in one way or another [20, 32–42]. For some techniques, additional textual [59] or geometric 3D [60–67] information is used. However, in terms of pottery research, the retrieval based on an incomplete query plays a significant role, since excavated objects are more often fragmented than not. This problem, commonly known as partial retrieval, is substantially more complex and is far from being solved for antique pottery [68]. At the same time, the necessary information is certainly present since experts are able to accurately guess the appearance of a whole vessel based on a single surviving sherd, using on both shape information of the fracture line and surface painting [64].

We propose the following solution. Instead of using solely a fragment as a query, we also incorporate expert knowledge regarding the shape in the query formulation. Sketching has been proven to be an intuitive way to introduce this kind of information [69–71]. We present a concept which makes use of the textual and shape information of a fragment, combined with expert knowledge introduced through a rough sketch outlining the missing geometry. This novel query modality (3D+Sketch) [72, 73] presents an intuitive and uncomplicated way for a user to specify missing shape information.

RQ2: *“How can a retrieval approach effectively support the search for specific motifs and patterns, exhibited by the vase paintings?”*

While a considerable amount of work has been done regarding the retrieval of pottery based on shape, the retrieval of vases based on their texture has received less attention [55]. Even

though some success could be achieved with off-the-shelf image retrieval techniques, an effective retrieval system for vase paintings needs to address several domain related challenges. First, when working with photographs of pottery objects, the paintings exhibit different degrees of non-linear distortions since they are on top of curved surfaces. The same is the case for renderings of 3D models. Hence, a reasonable preprocessing step of a texture-based retrieval system is to relax such distortions as good as possible. For 3D models, such can be achieved with projection techniques [74–77], while images can be relaxed similarly if the underlying geometry of the displayed object can be approximated [78]. Similarly to the shape-based retrieval, a challenge is the multimodality of data. Even if 3D surfaces are projected to a plane, which makes them comparable to photographs, a dataset can comprise colorized as well as grayscale images or even drawings. A typical way to bridge this modality gap is to bring everything into a common base modality, *e.g.*, transform photographs such that have the visual appearance of sketches [79] when drawings are compared to photographs. Even then, the comparison of entities is not straightforward, as the notion of similarity of many image-based descriptors per-default does not coincide with an archaeologist’s notion of similarity. The latter can be based on a motif’s silhouette (shape), its colorization (different clays), but also the context it appears in (combination with motifs/patterns). Hence, the definition of similarity is dependent on the concrete use case.

To address this complexity, we split the problem of texture-based retrieval into the problem of motif-based (humanoid or mythological figures) and ornament-based (all other surface paintings, *e.g.*, background fillings and such) retrieval. For the prior, we present an approach which allows an effective retrieval based on a motif’s silhouette [80]. The latter turned out to be a much more complex problem, as the ornamentation of surface paintings cannot be easily segmented into their atomic components. However, we propose a probability-based self-similarity detection which is able to recognize similar ornaments with minimal user intervention [81]. Moreover, we present an efficient annotation process for Peruvian [82] and Greek (Geometric) [83] pottery, which will provide a valuable starting point for future efforts towards further automation. For the prior, we also provide an initial benchmark dataset with 82 annotated objects, which is publicly available.

RQ3: *“How can interactive visualizations support an expert user in discovering (new) inter-object cross-correlations and links?”*

The high potential of interactive visualizations of large sets of objects has also been recognized in the CH community [21]. Different object traits are interesting for the computation of similarity clusters. Even though the subject of high-dimensional data analysis and visualization has been studied extensively [84], with techniques such as dimensionality reduction, projection schemes, parallel coordinates or specific interactive lenses, they generally assume that each object can be described by a real-valued vector. A problem with CH objects is that only some of their traits can be meaningfully expressed by a single scalar value, but most of them are either categorical (painting style, culture, epoch), or abstract (*e.g.*, shape, ornamentation). Hence, the majority of

established data visualization techniques are not directly applicable. Some approaches have been presented, specifically for CH applications, that combine two traits into an integrated visualization system [44, 50], but they are tailored explicitly to certain use cases and cannot be extended to support an arbitrary number of traits. Yet, at the same times, some correlations could only be unveiled through the consecutive analysis of several traits.

We present a scalable interactive exploration system [56] which supports that via a linked view concept. Correlations can be revealed by means of alluvial flows in-between views, which are displayed based on a selection or by an automatically determined correlation measure.

RQ4: *“How can the generation grammar, intrinsic to repetitive ornaments in vase paintings, be leveraged for the detection and the reconstruction of missing sequence parts?”*

Several works have been presented regarding the reconstruction of CH objects [6, 22]. In terms of ceramics, many completion techniques leverage the rotational symmetry typical for such objects [23–25], while some ‘hole filling’ concepts [26–28] try to fill holes based on the local context. All these methods have in common that they deal with the geometric completion of artifacts, while the textual completion seems to have received less attention [29–31].

We make a first step towards closing this research gap by proposing an approach for the automatic textual completion on 3D vessels [58]. More specifically, we investigate vessels exhibiting repetitive ornaments with some degree of arrangement (alternating pairs, checkerboard arrangements, *etc.*). By determining their generation grammar, we can detect gaps in a sequence and can present a hypothesis of what the missing texture parts could look like.

1.3 Scope of the Thesis

The scientific groundwork for this thesis was conducted from December 2018 till May 2022 at the Institute of Computer Graphics and Knowledge Visualisation (CGV) at the Graz University of Technology within the project “Crossmodal Search and Visual Exploration of 3D Cultural Heritage Objects” (CrossSAVE-CH¹, SEC. 1.3.1). The broad spectrum of topics resulted in a total of 13 peer-reviewed publications [42, 55, 56, 58, 72, 73, 78, 80–83, 85, 86] submitted to relevant international workshops and journals. The core of the thesis is based on 7 of them [56, 58, 73, 80–83] and comprises contributions to the fields of *Computer Vision* and *Visualization* but also *Applied Informatics*, *i.e.*, the application and evaluation of methodologies for a specific domain (CH in our case). Within the field of CH, the research was focused on ancient pottery of classical Greek and Peruvian societies. The decision to focus on this kind of objects results from the availability of data (which is abundant for pottery) as well as the existence of relevant use cases regarding search and exploration. That is, the researched methods and techniques are purely ‘object-based’, meaning that things like the reconstruction of CH sites, analysis of GIS

¹<https://www.tugraz.at/institute/cgv/research/projects/crosssave-ch/>

data, geophysical methods, *etc.* are beyond the scope of this thesis. Also, the conducted research was experimental in the sense that the development of a shippable software tool for everyday use was no goal of the project. Extensive usability testing with probands from without the project environment remains a future work item. Nonetheless, several prototypes and tools have been developed, some of which are publicly available², that can be used by other researchers for future endeavors.

1.3.1 The CrossSAVE-CH Project

The CrossSAVE-CH project was a three year interdisciplinary research project in collaboration with the *Institute of Archaeology* at the University of Graz and the research group for Pattern Recognition, Indexing and Social Media Analysis (PRISMA) at the University of Chile, funded by the Austrian Science Fund (FWF), grant no. P31317-NBL. The main research focus of this project is the searching, comparing, and visual exploring of 3D CH objects. A particular focus was on cross-modality with queries, including 3D models, images and sketches, and search spaces composed of 3D models, images, documents, drawings and metadata. Because of the oftentimes incomplete nature of CH objects, partial retrieval methods should be researched with the ultimate idea that a user can complete an incomplete 3D query with a sketch. Another focus of the project was the development of novel collection visualization techniques, as well as their evaluation using relevant archaeological 3D object and document data.

1.3.2 Research Approach

Within the first half of the project runtime, the research focus was on 3D object retrieval with an incomplete-3D-object-plus-sketch hybrid as a query. We initiated the project work by surveying related work pertaining to the topics 3D object retrieval [87], cross-modal retrieval [5, 20, 33, 36, 37, 88–93], and feature descriptors for 3D objects [60] as well as images [94, 95]. At the same time, we worked at building a multimodal pottery database for experimentation, which comprised a low number of 3D models provided by different national museums, as well as a significantly higher number of photographs which we obtained from online repositories and printed publications. From the data situation, it was clear that the primary data modality of the search space will be photographs. Hence, our design concept was inspired by established view-based retrieval approaches [88], which transfer a 3D query object into image space using rendering. We followed a similar design principle for a first prototype implementation in order to bridge the modality gap between the query and search space. With the query and database having the same modality, an important decision was the selection of an appropriate feature descriptor for comparison. To this end, we experimented with several color [96], texture [97], gradient [94] and shape-based [98, 99] feature descriptors, as well as learned image features. For the latter it was established that

²<http://datasets.cg.v.tugraz.at/pattern-benchmark/>

the features obtained with off-the-shelf image networks [100] are not sufficiently descriptive for the data at hand, while a from-scratch training of a network or transfer learning was not applicable due to a lack of available training data. From the compared concepts, a mixture of shape and color-based features yielded the most promising results. For evaluation purposes, we rely on synthetically fractured query objects, following the idea of Banterle *et al.* [91], who also use a Voronoi fracturing of inputs to this end.

During the experiments on the sketch plus 3D query hybrid, the relevance of textual information – in particular the search for semantically similar scenes – for the retrieval of pottery artifacts was recognized. A survey of the related work in this regard showed that the automatic semantic analysis of depicted scenes on vase paintings has not received much attention to date. As a starting point, we oriented ourselves by object detection techniques used in other domains. In the absence of any training data required for deep-learning based methods, it was established that a combination of both semantic segmentation [101, 102] and a segment-wise similarity measure based on shape and color [99] yielded the best results. While this concept proved to be efficient for the similarity retrieval of depicted living beings (humans and mythological creatures), which are commonly featured by Attic pottery (SEC. 1.1), very simplistic ornaments which are the atomic building blocks of Geometric era vase paintings or appear as background filling on Attic pottery turned out to be a limitation for the semantic segmentation as they are frequently intersecting each others. Yet, for this kind of problem, a sliding window confidence computation together with a Non-Maximum Suppression (two established techniques from the field of object detection) were able to provide promising results for several use cases. However, in general, the problem appeared to be far from being solved and most likely not being solvable with conventional methods. Hence, our focus shifted towards the application of deep learning techniques. To address the issue of unavailable training data, we compiled a set of labeled training data on ourselves, which involved the development of specialized annotation tools. The annotation collection was published, together with these tools and an evaluation metric, since similar efforts [103–106] have shown that the availability of benchmark datasets fosters the ambition of fellow researchers in finding solutions for a given problem. In the following we also experimented with recognition and classification based on CNNs by focusing exclusively on Geometric ornaments, a ‘smaller’ problem with a fixed number of distinct pattern classes given in a hierarchic taxonomy [17].

Interactive visualization was a focal point in the second project half. To this end, existing concepts [43–51] for visualizing large sets of CH artifacts were investigated. Especially influential was the 2019 survey of Windhager *et al.* [107] who reflect on the state-of-the-art of techniques for the visualization of digital CH collections. From that, it was determined that there is a research gap pertaining to the concurrent visualization of collections along an arbitrary number and combination of attributes. We developed a concept to this end, following the design triangle model by Miksch and Aigner [108]. The main idea behind this model is to consider *data*, *users*, and *tasks* which determine the visual representation and interaction methods for given analysis scenarios.

With a web-based prototype implementation, we evaluated several predefined archaeological tasks.

A final concluding research objective, we focused on towards the close of the project, is the topic of reconstruction. Even though this topic was not within the scope of the original project description, we found that the annotated patterns resulting from the benchmark dataset can be readily used to discover missing pattern entities. While many approaches have been presented for the geometric completion of CH artifacts [6, 23–28], the textual completion of repetitive surface patterns appeared to be a novelty. Preliminary experiments in this direction indicated that a detection of a pattern sequence’s inherent grammar can be used to detect gaps in the sequence, while a texture hypothesis for these regions can be generated using a machine-learning-based statistical model.

Over the whole project duration, the relevance of concepts for archaeological use cases was revised on a regular basis, with reoccurring discussions and design reviews with the archaeological partners from the University of Graz. Part of these meetings were the interaction with experimental prototypes. Those were developed following loosely the linear waterfall software development process by Royce [109], including the phases ‘Requirements’, ‘Analysis’, ‘Design’, ‘Coding’, ‘Testing’, ‘Operations’. The last two phases did, however, play an insignificant role in our case, since the development of shippable software was beyond the scope of the project.

1.3.3 Collaboration Statement

Conducting the necessary scientific work for writing up a thesis all by oneself is an insurmountable task, but many people are involved in the process of collecting ideas, reviewing relevant related work, revising paper drafts, *etc.* In short, all the different tasks pertaining to the creation of scientific papers. Stated in the following are all colleagues and collaborators who contributed to the work of this thesis in one way or another. This thesis would not exist without them.

Prof. Dr. Tobias Schreck from the Graz University of Technology is the supervisor of this thesis. He is head of the CGV, the institute where this thesis was written, and also project lead of the CrossSAVE-CH project, which constitutes the thematic framework for the presented work. He contributed with the conceptualization and the writing of all papers [42, 56, 58, 72, 73, 80–83, 85, 86] which emerged within the project and established valuable research connections to colleagues from the University of Graz as well as the University of Chile, resulting in several successful paper projects.

Dr. Reinhold Preiner from the Graz University of Technology is a research member of the CrossSAVE-CH project. As such, he participated in the writing of all project papers [42, 55, 56, 58, 72, 73, 78, 80–83, 85, 86], specifically by reviewing the technical soundness and correctness of formulas. In some cases, he also contributed with some of the necessary implementations [78, 80] or provided visualizations [55, 73, 78].

Dr. Stephan Karl from the University of Graz is specialized in Geometric, Corinthian, and Eastern Greek ceramics as well as 3D documentation and data analysis. He was the advisor from the archaeological side and helped in determining the relevance of applications for archaeological use cases, as well as with the selection of relevant scenarios for testing our proposed techniques. He authored [55] and co-authored [56, 58, 72, 73, 78, 80, 81, 83, 86] the majority of papers which emerged as part of the CrossSAVE-CH project and was mostly responsible for writing archaeology related text passages and checking the soundness of related statements.

Dr. Elisabeth Trinkl from the University of Graz is specialized in Greek Archaeology, Digital Humanities and antique textiles. As project member and archaeological advisor, she participated in the writing and revising of the majority of published papers [55, 56, 58, 72, 73, 78, 80, 81, 83, 86].

Prof. Dr. Ivan Sipiran from the University of Chile is a researcher from the field of computer science with a focus on 3D similarity search, geometry processing, computer vision and graphics in archaeology. He participated in the conceptualization and writing of several papers [58, 81–83] and also participated in the implementation in some cases [82, 83]. Furthermore, he is the initiator of the 2021 SHREC track [42] to which a solution was submitted.

Prof. Dr. Benjamin Bustos from the University of Chile is the primary research contact to the Millenium Institute for Foundational Research on Data and as such he was heavily involved in the defining the overall research direction as well as writing and revising the majority of published papers [42, 58, 72, 73, 80, 82, 83].

Dipl.-Ing. Peter Houska is a fellow PhD researcher from the Graz University of Technology and was involved in several fruitful paper collaborations [55, 78, 83] where he participated in equal part with conceptualization, implementation, evaluation and writing.

Dipl.-Ing. Alexander Komar was a research assistant for the CrossSAVE-CH project from 2018 to 2020. As such, he helped with all tasks pertaining the project work, including all sorts of implementation and evaluation, but also paper writing, resulting in him co-authoring several project papers [56, 72, 73, 80, 81].

MSc Arniel Labrada from the University of Chile is a fellow PhD researcher who collaborated on some project related efforts. In particular in experiments pertaining the applicability of deep learning techniques for given research questions. He also helped in the writing of three papers [72, 73, 80], where he focused on the related work sections covering relevant deep learning approaches.

Besides the above mentioned, several others contributed either directly or indirectly. That is, many of the developed concepts have been the result of feedback and discussion with various colleagues in the ongoing research seminars at the CGV while valuable input has also been provided by Prof. Dr. Kai Hormann from the Università della Svizzera italiana, Prof. Dr. Niloy J. Mitra from the University College London, and Dr. Paolo Cignoni from the Istituto di Scienza e Tecnologie dell'Informazione within the Doctoral Consortium at Eurographics 2019³, the 40th Annual Conference of the European Association for Computer Graphics.

1.4 Structure of the thesis

The current chapter is followed by two introductory chapters before CHAPTER 4–8 comprise the core technical contributions. Each of those chapters is based on one or more of our publications, which are outlined in a gray box at the beginning of a chapter. Due to the broad range of topics covered in this thesis, related work as well as a delineation of our contribution is stated in the individual technical chapters. They are grouped into four parts corresponding to the four main topics ‘Shape-based Retrieval’ (PART I), ‘Texture-based Retrieval’ (PART II), ‘Visualization’ (PART III), and ‘Digital Restoration’ (PART IV).

In short, the chapters contain the following:

CHAPTER 2 provides references to all the sources used to obtain the CH used in our experiments while also reflecting upon the different data modalities and describing how some types, *e.g.*, 3D models, can be obtained.

CHAPTER 3 describes the preprocessing steps to which some types of raw input data are typically subjected to. The chapter outlines different methods for flattening 3D shapes and gives an overview of image segmentation techniques, as well as their application to the data at hand.

CHAPTER 4 discusses the CBIR based on 3D queries. To this end, two orthogonal approaches are presented. First, a purely 3D based method [42] where a 3D model of a vase is compared against a database of 3D models with a customized descriptor. Second, the novel 3D+Sketch method [72, 73] which also supports cross-modal and partial retrieval efforts.

CHAPTER 5 presents a method [80] for the retrieval of typical scenes (warrior departure, wedding scenes) depicted in vase paintings by means of a query-by-example concept.

CHAPTER 6 is concerned with the ornamental information of vase paintings – *i.e.*, all the components of vase paintings which do not qualify as motif or background. In this regard, a method for the automatic recognition of such ornaments [81] is presented alongside interactive annotation concepts [82, 83].

³<https://www.eurographics2019.it/index.php/for-submitters/doctoral-consortium/>

CHAPTER 7 is focused on visualization and presents our concept of an interactive exploration system [56] for multimodal data.

CHAPTER 8 covers the topic of digital restoration with a particular focus on the topic of textual restoration, for which a novel method [58] is discussed.

CHAPTER 9 closes the thesis with a summary of all contributions, as well as a discussion of limitations and open research questions.

The convention for mathematical notations is outlined in APPENDIX A. Note that the scope of used symbols is limited to the chapter they appear in.

DATA SOURCES AND ACQUISITION

Contents

2.1	Data Sources	18
2.1.1	Repositories	18
2.1.2	Print Volumes	19
2.2	Data Modalities	20
2.2.1	Texts	21
2.2.2	Metadata	21
2.2.3	Profile Curves/Archaeological Drawings	21
2.2.4	Photographs	23
2.2.5	3D Models	24
2.3	Acquisition	24
2.3.1	Data Extraction from Archaeological Volumes	26
2.3.2	Crawling of Web Resources	28

Relevant digitized data can be obtained from various sources and is given in different data modalities (*c.f.* FIG. 1.2). In domain publications, for example, the information on a specific object is commonly given as a set of photographs together with textual descriptions and – in rare cases – archaeological drawings of the vase paintings. SEC. 2.1 presents the different data sources used to obtain the relevant object information necessary for experiments, while SEC. 2.2 discusses in detail the modalities they appear in. Finally, SEC. 2.3 concludes the chapter with a discussion of the techniques used for the acquisition, *i.e.*, the processes involved in transforming the raw data into structured object records, such as the segmentation of printed pages or the crawling web resources.

2.1 Data Sources

The long-time preservation of scientific data is an essential objective for all scientific fields. For our purposes, the most easily accessible data source is online repositories, as they provide data already in a machine-readable and structured manner. Also, suchlike provided data usually does not exhibit any digitization artifacts, such as, *e.g.*, rasterisation artifacts in cases where printed photographs are scanned or erroneous translations if images of texts are converted into a machine-readable representation. However, on the downside, much pottery-related data is only available in printed publications, with the original digital copy not being released by the publisher. As this is the only form of documentation for a large part of excavated artifacts, this type of source cannot be ignored. Yet another source of data is the peer-to-peer exchange with fellow researchers. Within our close cooperation with archaeologists working in the field of ancient pottery and its digitalization, we have access to several sets of scanned 3D objects provided by the courtesy of several national museums.

2.1.1 Repositories

Our work's most pivotal web repositories are the already mentioned BAPD and the CVA Online. The CVA Online [18] is an open-access digital element of the CVA project (*c.f.* SEC. 2.1.2), which offers several of its print fascicules in a digitized manner. It started with approximately 250 fascicules in 2000 and contains today more than 350. Digitized text pages and plates (the pages in the second half of a volume containing the photographic information) can be reviewed and browsed through the site. Illustrations and other selected content have also been merged with the BAPD.

The BAPD is the world's largest repository of ancient Greek pottery, featuring almost 130,000 records with about 250,000 images. It originates from the physical Beazley Archive, located in Oxford's Classical Art Research Centre. The data has been collected from museums, auction houses, and digitized volumes of the CVA series and comprises mostly pots produced from the 6th to 4th century BCE in Athens [8].

There are also many other repositories, mostly curated by museums, a selection of which is listed in TAB. 2.1. Besides references, the table also shows the data modalities featured by the respective collections. Even though this list is by no means encompassing nor extensive – additional online sources are stated by Kurtz *et al.* [40] and Scopigno *et al.* [52] – we can observe that almost all repositories comprise at least metadata and photographs. Textual descriptions and 3D data appear to play a less significant role. We attribute this to the fact that the display of 3D models in a browser is substantially more complex than the display of photographs, since it usually requires more data to be transferred. Also, not yet all browsers and platforms support 3D viewers.

TABLE 2.1. A small selection of online data sources from different museums and research facilities, as well as the data modalities used for documentation: text, metadata, drawing, photograph, and 3D model.

Resource	Text	Metadata	Drawing	Photograph	3D
CVA Online [18]	●	○	●	●	○
BAPD [8]	○	●	●	●	○
Athenian Agora Excavations, American School at Athens [110]	○	●	●	●	○
Corinth Excavations, American School at Athens [111]	○	●	●	●	○
British Museum [112]	○	●	○	●	○
Museum of Fine Arts, Boston [113]	●	●	○	●	○
Ure Museum of Greek Archaeology, University of Reading [114]	●	●	○	●	○
The J. Paul Getty Museum Collection [115]	○	●	○	●	○
Ashmolean Museum, Oxford [116]	○	●	○	●	○
Digital Archaeological Archive of Comparative Slavery [117]	●	●	●	●	○
Europeana [118]	○	●	●	●	●
3D Kulturdatenbank, Landessammlung Niederösterreich [119]	○	●	○	●	●
The Virtual Hampson Museum [120]	●	●	○	●	●
Sketchfab [†] [121]	○	○	○	○	●

[†] Sketchfab is not dedicated to the hosting of curated CH data but any 3D data, which can be uploaded by users. However, it nonetheless features a large amount of high quality 3D CH objects – some of which are artistic designs, but others are valid models scanned from real artifacts.

2.1.2 Print Volumes

The most crucial resource for high-resolution photographic data is the fascicules of the CVA project, a research project by the Union Académique Internationale, started in 1922. While there are dedicated publications documenting individual CH artifacts in depth [74], the CVA fascicules provide thematically grouped collections in a reasonably structured manner and follow established conventions, thus allowing the application of automatic data extraction techniques (SEC. 2.3.1.1). To date, more than 400 fascicules from 24 countries have been published, documenting more than 100,000 vases. Even though most of those fascicules are also freely accessible in the CVA Online, the quality of the scanned pages, most of which were obtained in the early 2000s, is relatively poor. *I.e.*, most pages exhibit visible grain and even artifacts of paper texture and are thus only marginally usable for image-based retrieval methods. Also, while irrelevant for a human user, the presence of watermarks compromises unsupervised methods for data extraction and feature computation. Consequently, we mainly relied on a set of fascicules available to us in physical form and previously scanned by our project partners. TAB. 2.2 shows our selection of CVA volumes for photographic data. In total, we extracted 3,340 photographs belonging to 1,173 separate exhibits. We focused on Geometric pottery and selected especially recent installments, as they exhibit a generally superior photo quality.

TABLE 2.2. Delineation of photographs extracted from different CVA fascicules and the BAPD. In cases where data for a fascicule was extracted from both sources, the object records were merged using the plate and figure number of a photograph as a unique identifier.

Resource	Year	Photographs	Objects	Source	
				CVA	BAPD
CVA Karlsruhe 1 [†] [122]	1951	19	15	●	○
CVA München 3 [†] [123]	1952	139	68	●	○
CVA Kassel 1 [†] [124]	1972	363	146	●	●
CVA Louvre 16 [125]	1972	162	54	●	○
CVA Würzburg 1 [†] [126]	1975	92	38	●	○
CVA Tübingen 2 [†] [127]	1979	87	39	●	○
CVA Athens 5 [128]	2002	302	81	●	○
CVA Metropolitan Museum 5 [129]	2004	277	112	●	○
CVA Erlangen 2 [130]	2007	534	162	○	●
CVA Göttingen 3 [131]	2007	672	200	○	●
CVA Berlin 10 [132]	2009	645	213	●	●
CVA München 16 [133]	2010	556	55	○	●
CVA Oxford 4 [134]	2010	363	168	●	○
CVA British Museum 11 [135]	2010	363	206	●	○
CVA Jena 1 [136]	2011	479	109	○	●
CVA Berlin 13 [137]	2013	386	62	●	●
CVA Dresden 2 [138]	2015	558	88	○	●
CVA Bonn 4 [139]	2016	276	93	●	○
CVA Berlin 17 [140]	2018	595	113	●	○

[†] Only plates containing geometric pottery have been extracted.

Another print volume that we used for building a taxonomy of Geometric patterns is the *Ornaments of Geometric Vessels* (in German) by Kunisch [17]. The volume comprises a total of 776 distinct pattern classes collected from object depictions from more than 50 different domain publications. The entries in this publication were also extracted automatically (SEC. 2.3.1.2).

2.2 Data Modalities

In the following, the data modalities relevant to the presented methodologies – text, metadata, drawing, photographs (gray, colored), and 3D models – are briefly described. This list is by no means encompassing since other modalities, such as video, sound, and 1D data (time series, digitized voice, or music), exist. However, they play an insignificant role in archaeological documentation and are beyond the scope of this thesis. The commonness of different data modalities for CH documentation is also discussed by Windhager *et al.* [107].

2.2.1 Texts

Texts were primarily extracted from CVA fascicules. Usually, the running text sections in a fascicule's first half are structured so that one paragraph refers to one plate. Textual information can contain, but is not limited to, where similar objects are published, characteristics of a vessel, and (if it is present) a description of the vase painting. Generally, these descriptions are non-standardized and unstructured, describing either the object as a whole or referencing certain parts, *e.g.*, a vessel's shoulder. The texts also contain information on which photographs belong to which object, allowing to establish image-to-object relations automatically. The descriptions lack, however, any direct reference to specific regions of a photograph. A challenge inherent to the CVA fascicules is that texts are in the official language of the country where it is published. Hence, they comprise many languages, complicating the problem of automatic processing and computing text similarities.

2.2.2 Metadata

The objects featured in the BAPD can contain up to 13 (optional) entries on an item's context, custodianship, and coverage in existing publications, among others. According to Orton [141], the four most essential object properties for pottery research are date, distribution, functionality, and state of preservation. Martínez-Carrillo [59] similarly states the importance of the spatial and temporal context for ceramic analysis. The structured metadata used within the scope of this thesis was obtained from the BAPD, where it can be easily extracted from the HTML structure of object records. Different properties are presented as key-value pairs and often contain the fields (i) 'Fabric', an object's cultural context, (ii) 'Technique', the painting style such as black figure or red figure, (iii) 'Provenance', an object's provenance, (iv) 'Date', the assumed date of production, (v) 'Decoration', a short description of the vase painting, and (vi) 'Publication Record', the references to publications the object is featured in.

2.2.3 Profile Curves/Archaeological Drawings

Profile curves are an obviously well-suited modality for documenting a vase's shape due to the rotational symmetry of such wheel-manufactured objects. Hence, they have been the method of choice to this end throughout the history of pottery research. In archaeological publications, such formal descriptions of ceramic shapes usually consist of a bipartite artistic representation of a vase, where one half depicts the artifact as a solid object (sometimes with additional cues regarding the vase painting), while the other half shows a slice through the rotational solid, illustrating the different wall thicknesses along the profile curve (FIG. 2.1). If the vase exhibits appendages like handles, cross-sections of them as well as codes referring to their type, number, and position, can accompany the profile drawing [141, p.192-193]. Orton [141, p.196] describes the process of manually obtaining a profile curve as follows: "The vessel profiles are first digitised

using a digitising tablet and the data standardised to the same height, thus eliminating the overall size as a factor and enabling researchers to concentrate purely on the shape element.” For varied shapes, the procedure is repeated at different radial slices to account for minor local variations before the interpolation between these results is taken as the final profile curve. If a 3D model of a vase is available, the profile can also be determined automatically (see, *e.g.*, SEC. 4.3). However, experiments by Stephan Karl (FIG. 2.1) showed that the increased accuracy of profile curves obtained with digital methods is beyond what is relevant for pottery research. *I.e.*, profile drawings are not only the most frequently used method for shape description in historical documentation, but also constitute the state-of-the-art.

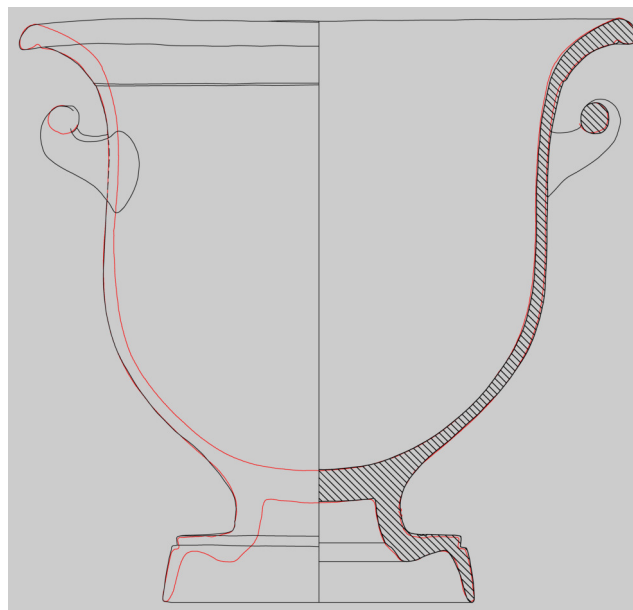


FIGURE 2.1. Most drawings of vase shapes combine a representation of the solid of revolution (left-hand side) with optional details about the vase painting and a slice-through (right-hand side), illustrating the various wall thicknesses. The overlaid red outlines constitute the profile cross-section obtained from 3D data. (*Image source: Stephan Karl*)

The creation of scientific drawings of vase paintings is an essential practice in archaeological research and documentation. Over time, the nature of such drawings changed alongside their intended application. While iconographic studies were the focus of initial research efforts, more modern approaches pioneered by the Italian Giovanni Morelli focus on classifying individual artists and workshops [10, 142]. This also resulted in a shift in the creation of archaeological drawings, which were formerly interspersed with borrowed fashions, idiosyncratic habits, and mistakes, towards more precise representations (*e.g.*, FIG. 6.17). Traditional approaches for the creation of such drawings include manual drawings using pen and paper, either directly on the physical surface using tracing paper or from photographs [142]. Modern approaches also make use of 3D digitized artifacts, which allow contactless analysis and documentation using so-called unwrapping techniques [55].

2.2.4 Photographs

Photographs are the data modality with the vastest amount of available data since they are comparably effortless to produce and have been a means of documentation for decades. The photographs used in our experiments are obtained from two different sources – the CVA and the BAPD. From the CVA, we used 19 fascicules, illustrating a wide range of different vessel types in color as well as grayscale. TAB. 2.2 outlines a breakdown of the number of photographs obtained from different sources and the number of objects they can be attributed to (as one object is typically represented with more than one photograph).

Unfortunately, the relation of photographs to individual objects is non-trivial to extract automatically, as this information can only be learned from the relatively unstructured running text. However, a photograph is unambiguously identified by its plate and figure number. Hence, we can leverage our second source of information, the BAPD, which provides object-based records that contain the information where an object is published in a (semi-)standardized manner. *E.g.*, the string “Corpus Vasorum Antiquorum: BERLIN, ANTIKENSAMMLUNG 13, 53-54, BEILAGE 10.1, PL.(4531) 38.1-5”¹ denotes that the object in question can be found in CVA Berlin 13 [137] and is represented with figures 1 to 5 on plate 38. Based on that, we can automatically piece together image-to-object relations. FIG. 2.2 shows a randomly selected sample of the so compiled database. Note that it features not only photographs depicting objects from canonical views,



FIGURE 2.2. A small sample of the search space comprising almost 7,000 photographs. Besides depictions from canonical views, they also comprise closeups of details/motifs or photographs of sherds.

but also several close-ups of specific details or photographs of sherds. Additionally, 1,548 object records with a total of 3,471 photographs were extracted from the BAPD together with structured metadata.

¹www.beazley.ox.ac.uk/record/2E49939D-7550-4278-AC5F-9C9BE2941B04

There is also a potentially high number of tourist photographs of display cabinets available on different image-sharing platforms. For the scope of this thesis, however, such image sources have been omitted since they are not photographed in the standardized manner of archaeological photographs, *i.e.*, they often show multiple objects with distracting gradients in the background.

2.2.5 3D Models

Today, high-resolution textured 3D models can be obtained from contact-less acquisition methods like *Structure-from-Motion* [143] or *Structured Light Scanning* [144]. The representation of 3D shapes is a research field in its own right and includes concepts like *implicit surfaces*, *unstructured points* (point cloud), and polygonal meshes (*e.g.*, triangle meshes), among many others [145, 146]. The most common formats for archaeological pottery research are either high-resolution triangle meshes with vertex colors [74] or coarser meshes with texture maps [42]. Untextured meshes are also used for other CH content [119] if the texture is irrelevant.

The 3D models used within this thesis are provided by courtesy of the Karl Franzens Universität Graz (KFU), the Landesmuseum Kärnten (LMK), the Universalmuseum Joanneum (UMJ), and the Kunsthistorisches Museum Wien (KHM). They comprise high-resolution 3D scans with vertex colors, representing pottery from the Geometric period (c. 9th - 8th century BCE) to the Classical period (c. 5th - 4th century BCE). An extensive collection of pottery objects from pre-Columbian cultures in Peru with varied geometry and artistic styles, captured by the Museo de Arqueología Josefina Ramos de Cox (MAJRC) in Lima, Perú, has also been available as part of a recent shape retrieval contest [42]. The models, given as textured triangle meshes, were manually oriented such that the rotation axis is co-aligned with the *Y*-axis. Furthermore, all of them were simplified to consist of about 40,000 triangles.

TAB. 2.3 provides a listing of all available 3D objects by their source. All registered artifacts are assigned an inventory number, a combination of the abbreviated venue name combined with a sequential alphanumeric identifier, *e.g.*, *KHM IV 1* for the pitcher featured in FIG. 4.11. These numbers uniquely identify an object and are used throughout the thesis to refer to specific artifacts.

2.3 Acquisition

With ‘acquisition’, we do not refer to the capturing techniques used to obtain archaeological data. However, to this end, we point to respective archaeological literature for producing drawings [141], photographs [53, 54, 147], and 3D models [143, 144]. In contrast, this section discusses the steps necessary to extract textual and figurative information from (unstructured) documents and group these data items according to their affiliation to a particular object (SEC. 2.3.1). *I.e.*, the process of transferring raw data into a ‘per-object’ data structure. The situation is different for web repositories (SEC. 2.3.2), another rich source of information, which provides object information

TABLE 2.3. An overview of the different sources for the 3D models, used throughout the thesis.

Resource	Abbr.	Objects	Colorization	Example
Universalmuseum Joanneum	UMJ	10	Vertex Colors	
Landesmuseum Kärnten	LMK	17	Vertex Colors	
Kunsthistorisches Museum Wien	KHM	11	Vertex Colors	
Karl Franzens Universität Graz	KFU	10	Vertex Colors	
Museo de Arqueología Josefina Ramos de Cox	MAJRC	938	Texture Maps	
3D Kulturdatenbank Landessammlungen Niederösterreich	-	53	-	

usually in an already structured manner. In this case, tools and libraries can be used to acquire portions of a database automatically.

2.3.1 Data Extraction from Archaeological Volumes

In the following, we discuss automated workflows for extracting the figural information contained in CVA fascicules as well as all the information pertaining to Geometric patterns from *Ornaments of Geometric Vessels* [17]. To this end, we employ basic computer vision techniques combined with a set of simple heuristics.

2.3.1.1 Plate Segmentation of CVA Fascicles

The figural data in the CVA fascicules is contained within the last half of the documents on dedicated pages, referred to as ‘plates’. To obtain the individual photographs of such a plate, we use a *segmentation* (more on that in SEC. 3.2) approach. The idea behind this process is to transfer the image of a plate into a binarized representation, where connected components correspond to the individual images. Initially, the raw plate image (FIG. 2.3(a)) is converted to grayscale and subjected to an intelligent histogram adjustment, which sets all pixels within the top- and bottom 4% of intensities to the respective maximum and minimum intensity values. That ensures that the pixels of the almost white plate background are located on the high end of the color histogram, permitting the application of a tight threshold for generating a binarized image. FIG. 2.3(b) shows an exemplary result of this step. We can see (red and blue closeups) that, in some cases image numberings or noise in the inputs account for sprays of white pixels between images, blending the components together. We counteract this unwanted behavior by applying an opening transformation with a small rectangular structuring element. A subsequent closing transformation with a slightly larger kernel size reunites components that belong together while still maintaining the desired gaps between components belonging to different images. Based on this result, the connected components with 8-way connectivity are determined. The bounding boxes around the component labels are used as extraction masks (FIG. 2.3(c)) after ruling out too small components whose bounding boxes take up less than 0.5% of the overall plate area.

In addition to the depictions themselves, we extract the associated figure number, which uniquely identifies them. In most cases, these numbers are located below and either centered or on the left-hand side of their associated image. A small image segment of the assumed location is used as input for an Optical Character Recognition (OCR). To this end, we apply the Tesseract Open Source OCR Engine [148]. If no string is found or the string does not fit the expected format of numerical characters, the segment is enlarged until a valid string is found or the segment exceeds a predefined threshold size relative to the plate size.

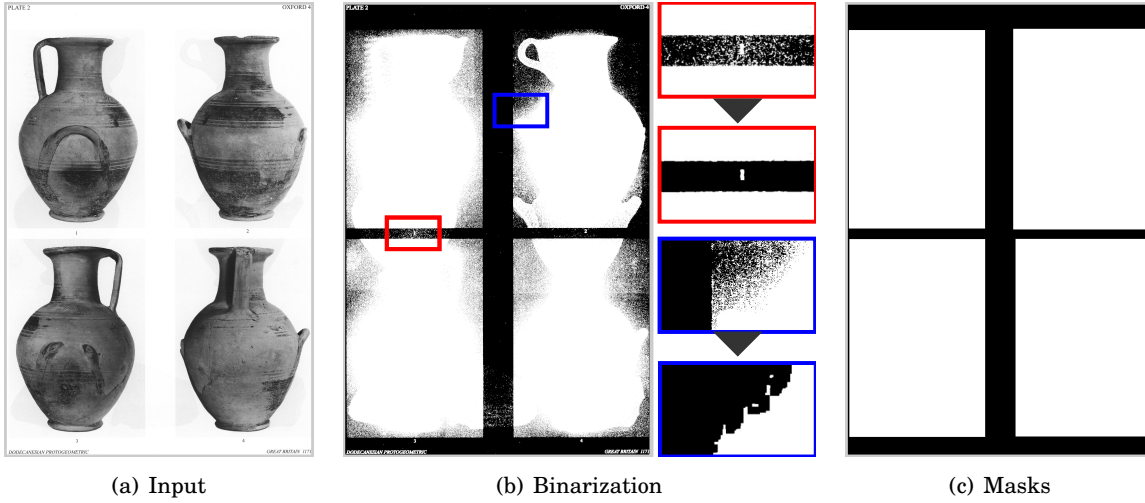


FIGURE 2.3. A typical plate [134, pl.2] (a) is binarized by thresholding (b). Noisy artifacts are removed by morphological opening (red section in (b)) and tiny components are merged with morphological closing (blue section in (c)). Bounding rectangles around detected contours (c) are used as extraction masks.

2.3.1.2 Pattern Taxonomy Extraction from Print Volume

The Geometric pattern taxonomy by Kunisch [17] comprises a total of 776 distinct pattern classes collected from object depictions from more than 50 different domain publications. They have been grouped into 6 major – e.g., *rectangular ornaments* – and 20 minor groups. A specific pattern class is generally documented by:

- A representative black and white drawing of the pattern, usually closely resembling one of the related examples on a real artifact;
- A textual description, consisting of different labels referring to certain properties of the respective pattern, like ‘hatched’, ‘vertical’, *etc.*, in up to five different languages; and
- (Optionally) literature references to publications depicting objects the pattern appears on.

This format is not entirely consistent, though, as several descriptions refer to multiple pattern drawings. As the book was not available in digital format, the contents were extracted from the scans of a printed version using basic image and text processing techniques. In the first step, all text depictions were purged from the binarized pages with a morphological erosion, leaving only the figures containing the pattern drawings (FIG. 2.4, blue). After cutting out all figures and flood-filling their respective rectangles with white, all remaining non-white areas were qualified as ‘text blocks’ (FIG. 2.4, yellow) and extracted as well. The figures were automatically segmented into individual pattern drawings (FIG. 2.4, red) by employing the connected component analysis method [149], implemented in the OpenCV computer vision library², together with simple

²<https://opencv.org>

heuristics. The same procedure was used to segment the text blocks into individual pattern descriptions (FIG. 2.4, green). The figure captions, combined with their respective subfigure captions, were leveraged to link the drawings with their respective descriptions. OCR [148] was used to transfer the rasterized text into a machine-readable form. All in all, about 90% of the entries were segmented, translated, and linked correctly with these methods, while the rest was fixed manually.

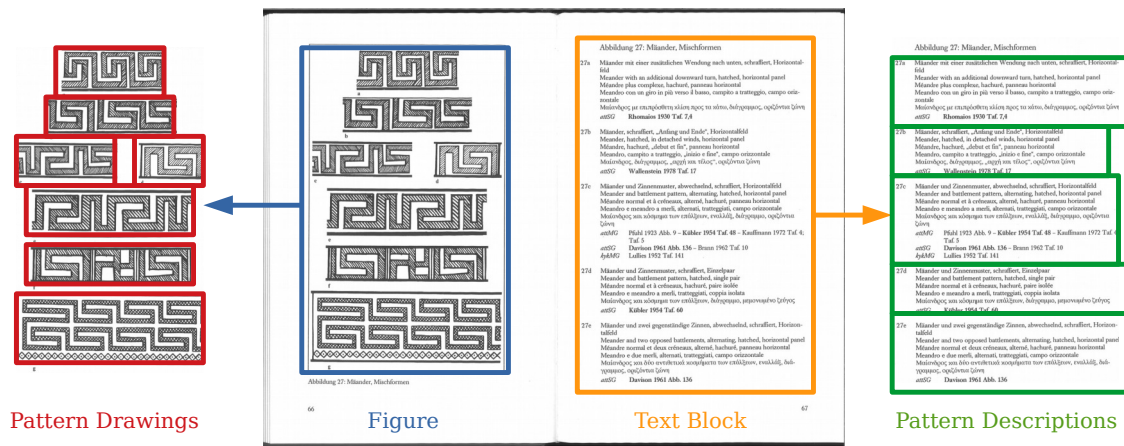


FIGURE 2.4. One characteristic double page from the geometric pattern taxonomy by Kunisch [17, p.66-67]. The figures (blue) and text blocks (yellow) were detected and extracted automatically and further segmented into pattern drawings (red) and descriptions (green), respectively.

2.3.2 Crawling of Web Resources

Web repositories are an ideal source of information, as they provide object information in an already structured manner. *I.e.*, most of them provide dedicated detail pages for individual object records comprising photographs, 3D models, free text descriptions, and structured metadata. The Document Object Model (DOM) of such an HTML page can be leveraged to automatically extract the displayed information and link it to specific object properties through a combination of parsing³ and a set of well-defined heuristics. Regular expressions applied to search results can be used to obtain the links to different detail pages, while dedicated software packages⁴ can be employed to retrieve them.

Within the scope of our project, a combination of such low-level web crawling techniques was used to compile a total of 1,424 object records with 6,955 images and associated metadata from the BAPD.

³<https://www.crummy.com/software/BeautifulSoup/>

⁴<https://www.gnu.org/software/wget/>

PREPARATION

Contents

3.1	Surface Unwrapping	30
3.2	Segmentation	31
3.2.1	Object Segmentation	32
3.2.2	Motif Segmentation	34

As indicated in FIG. 1.2, some data needs to be further processed before it can be used for specific domain applications. These preprocessing steps are varied and depend on the data modality. In terms of 3D data, this involves, for example, mesh cleaning actions like removing non-manifold edges, isolated pieces, degenerated faces, and self-intersecting faces, as well as closing holes and correcting normals. When used for retrieval, 3D models are usually subjected to several normalization steps, such as translation (to the origin), scaling, orientation along the principal axes, denoising, and flipping [87]. In pottery research, other domain-specific tasks like axis estimation, capacity calculation, feature segmentation, profile extraction, surface unwrapping, and pattern segmentation also play a significant role [55]. Segmentation is also relevant for photographic data. Either in the form of Foreground/Background (FG/BG) segmentation, distinguishing between pixels showing the object and pixels containing the background, but also in the form of semantic segmentation, breaking down the patterns and motifs exhibited by the object's surface.

In the following, we discuss two preprocessing tasks in-depth, which have been pivotal for our proposed methods: The unwrapping of 3D surfaces (SEC. 3.1) and the segmentation of archaeological photographs (SEC. 3.2).

3.1 Surface Unwrapping

The unwrapping of vessel surfaces is a long-standing practice in archaeology since such projections allow us to perceive an object's vase painting as a whole, as opposed to dividing them over multiple images exhibiting photographic distortions [142]. Unwrappings are not only a beneficial format for documentation, but also advantageous for annotation, retrieval, and automated analysis of vase paintings. However, the 'manual' unwrapping with tracing paper or the like (archaeological drawing) is very time-consuming, error-prone, and oftentimes neither possible nor allowed due to the fragile nature of CH artifacts. That is, if a 3D model is available, computerized methods can be applied to this end.

The unwrapping of 3D shapes is intensively studied in the field of computer graphics for applications like texture mapping or geometric modeling. Extensive surveys regarding these topics are given by Floater and Hormann [150] and Sheffer *et al.* [151]. While computer graphics is concerned with unwrapping arbitrary shapes, the problem can be substantially simplified for the pottery research domain since shapes can be generally regarded as revolutional solids. This circumstance promotes the applications of methods that approximate the shape by fitting a simple proxy geometry, *e.g.*, cylinders [76], (combinations of) cones [75, 152, 153], or spheres [74]. The development of the surface of these proxies is known and can thus be used to unwrap any given input. These techniques are implemented in the *GigaMesh*¹ software framework. However, such projections can lead to different types of unwanted distortions. Wang *et al.* [154] aim to minimize these distortions by segmenting the surface of the input shape in a regular arrangement of equal-sized tiles, which are then unfolded separately. Preiner *et al.* [77] solve this problem in an alternate fashion by globally relaxing the stresses induced by a conventional projection scheme using a physical mass-spring model. Later, Houska *et al.* [78] extended this method such that unwrappings can also be obtained from a sparse set of canonical photographs instead of a 3D model. To this end, they extract the object silhouettes from the individual photographs, which are then used to estimate the shape of the displayed geometry. Based on this estimate, each photograph is 'flattened' independently before they are stitched together, and distortions are minimized with an elastic relaxation.

Throughout this thesis, the unwrapping by Preiner *et al.* [77] and an adapted variant of the cylindrical unwrapping by Karras *et al.* [76] are used. Both methods have their advantages and disadvantages. *I.e.*, the elastic relaxation should have less overall distortion, yielding visually more pleasing results. On the other hand, it induces other non-linear distortions, *e.g.*, straight horizontal and vertical lines will no longer be straight in the relaxed output. For some automatic processing tasks, it is, however, beneficial to keep this constraint. In such cases, cylindrical unwrapping is more appropriate. Furthermore, it has the advantage that the method is more robust and can be computed more efficiently. We modified the cylindrical projection proposed by Karras *et al.* such that the proxy cylinder is non-linearly elongated in the direction of the rotation

¹<https://gigamesh.eu/>

axis relative to the arc length $\int \sqrt{1 + [r'(z)]^2} dz$ of the vessel's profile curve $r(z)$ (FIG. 3.1). Said results in less distortion in high-curvature areas such as the rim or the shoulder.



FIGURE 3.1. The 3D model of an Attic black-figured amphora (left) is unwrapped with the cylindrical unwrapping by Karras *et al.* [76] (middle) as well as with our adapted variant (right). The latter result in a visibly less distorted projection in the shoulder region. The unwrapping obtained from an elastic relaxation for the same vase can be seen in FIG. 6.20.

3.2 Segmentation

Generally, the process of segmentation splits a given input into distinct regions, representing meaningful parts. The topic represents a broad research area with various methods proposed over the years. Depending on the type of input, we can distinguish between the two major groups of *Image Segmentation* and *3D Object Segmentation*. Depending on the objective, image segmentation can be further classified into three subcategories: (i) *Semantic Segmentation*, the assignment of a class label to each pixel, *e.g.*, the class ‘people’; (ii) *Instance Segmentation*, the detection and segmentation of distinct object instances, *e.g.*, ‘person0’, ‘person1’, *etc.*; and (iii) the recently proposed *Panoptic Segmentation*, which represents a combination of the prior two [101].

Various semantic segmentation methods are mentioned in Gonzalez *et al.* [102], including the construction of region boundaries based on detected edges, or the growing of regions from seed points based on texture similarity or other local image properties. Also, morphological operations or fitting of pre-defined generative templates are known approaches. More recent techniques mostly rely on the application of deep learning. *I.e.*, Girshick *et al.* [155] and Sermanet *et al.* [156] applied CNNs to identify candidate regions of objects within an image. Other leading works in the field, which achieve state-of-the-art performances, use a Fully Convolutional Network (FCN) architecture [157, 158]. CNNs, originally developed for object detection, have also shown outstanding performance for the task of instance segmentation. *E.g.*, Mask Region-based CNN (R-CNN) [159] and Feature Pyramid Network (FPN) [160] are such adapted end-to-end deep neural network architectures. Likewise, many techniques exist for 3D object segmentation [161], ranging from simple partitioning schemes [162] to the derivation of 3D skeletons [163] or graph-based approaches [164, 165].

In general, there is no universal approach to segmentation, but it has to be guided by application requirements, specifying, *e.g.*, the level of detail at which segments are to be distinguished. Within the work for this thesis, segmentation was used for three different purposes: (i) detecting and extracting the individual photographs on CVA plates (SEC. 2.3.1.1), (ii) removing unnecessary background information from these photographs with a FG/BG segmentation (SEC. 3.2.1), and (iii) splitting the surfaces of painted pottery into cohesive areas corresponding to motif instances (SEC. 3.2.2). As we use segmentation for processing large datasets, we require methods that work entirely unsupervised and efficiently.

3.2.1 Object Segmentation

Even though scientific artifact photographs are taken under laboratory conditions with a neutral background and favorable lighting conditions, they usually include additional scene information, mainly in the form of shadows cast by the displayed object itself (FIG. 3.2(a)). Said is problematic for two reasons. First, the scene information introduces additional gradients, influencing gradient-based feature descriptors with unwanted information. Second, it prohibits the straightforward detection of the object contour, which is the basis for contour-based retrieval approaches.

This problem can be addressed with a semantic FG/BG segmentation, which allows to distinguish between object and background pixels and lets us consequently disregard all unwanted background information. We employ a segmentation based on the graph-cut extraction algorithm by Rother *et al.* [164]. The idea of this approach is to treat an image as an unconnected graph $G = (V, E)$ where the vertices V correspond to pixels and the edges E are used to model the neighborhood of pixels. The edges carry a weight relative to the dissimilarity between two pixels. The aim is to split the graph into connected components corresponding to the foreground and background pixels while maintaining minimal cutting energy. Graph-cut approaches are typically interactive, as initial seeds are needed to initialize the individual nodes as possible foreground or possible background. As we require a purely unsupervised solution, we generate the initial seed automatically based on several meaningful heuristics.

First, the input is subjected to the same histogram adjustment as the plate images (SEC. 2.3.1.1) to enhance the contrast between object and background pixels (FIG. 3.2(b)). A thresholding is applied to bring the image into a binarized representation. Archaeological photographs are in grayscale as well as in color. For grayscale images, this step is straightforward. For colored inputs, it has been found that the blue color channel exhibits the best contrast and has thus been used for thresholding. From the example in FIG. 3.2(c), it can be seen that thresholding alone does not necessarily result in a proper object mask. On the one hand, there are remnants of shadows in the background that introduce artifacts in the binarization. On the other hand, some regions on the object surface (especially those displaying motifs) are erroneously marked as background. To overcome the prior, a morphological closing operation with a small kernel size is applied (FIG. 3.2(d)). To address the latter, we take the convex hull of all remaining mask pixels

as possible foreground (FIG. 3.2(e)). The resulting mask is then used as the seed for the graph-cut algorithm, resulting in a proper object mask (FIG. 3.2(f)).

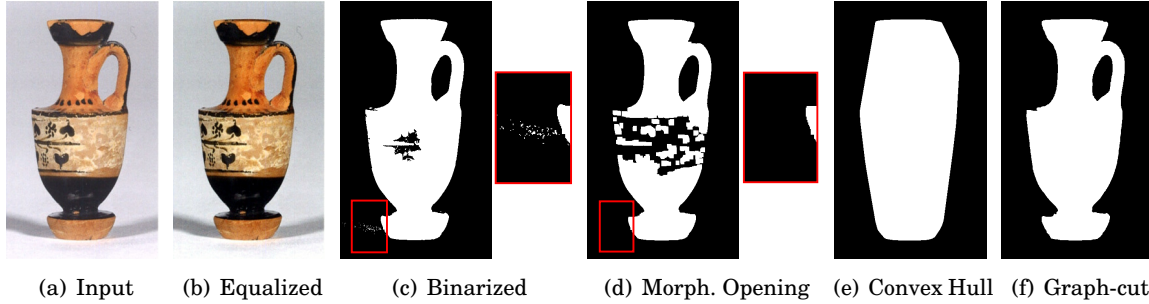


FIGURE 3.2. The raw input (a) is subjected to histogram equalization (b) and thresholding (c). With a morphological opening, relicts of the background scenery are removed (d). The convex hull of all remaining connected components (e) is used for a graph-cut extraction (f).

We assume that objects are generally photographed in a somewhat standardized manner [53, 54, 147], showing objects from canonical angles with a near orthographic projection in front of a bright neutral background. The hyperparameters for thresholding, morphological transformations, and the graph-cut were tuned on a small representative sample of the whole image database. Unfortunately, some photographs differ from the proposed standards and show, *e.g.*, an illuminated object in front of a dark background (FIG. 3.3, left) or feature prominent drop shadows resulting from illumination (FIG. 3.3, middle). In such cases, the object segmentation will not work reliably. Another limitation can be caused by the depicted objects themselves if parts of their surface exhibit a very bright colorization (FIG. 3.3, right).



FIGURE 3.3. Different edge cases limit the applicability of automatic object segmentation. Those include divergences from proposed standards like a dark background (left), prominent drop-shadows resulting from illuminations (middle), or object parts with a very bright colorization (right).

3.2.2 Motif Segmentation

The aim of the motif segmentation is to further split the extracted object depictions into cohesive areas corresponding to the outlines of the various motifs shown by the vase paintings. We refer to ‘motif’ in the sense as defined in SEC. 5.1. Hence, ornamental bands or simple shapes appearing as background fillers are deliberately omitted. While there are also numerous approaches for 3D segmentation, we focus exclusively on the extraction of motifs from photographs, the modality the vast majority of our data consists of. In contrast to the FG/BG segmentation used for masking the background, we need to employ an instance segmentation method, since multiple motif classes with multiple instances can be depicted on the same vase. A variety of approaches exists for this very purpose [55], all of them exhibiting different data-dependent strengths and weaknesses. Two orthogonal approaches have been applied and evaluated for the data at hand: a graph-based approach utilizing the gradient information of a grayscale image and an approach based on morphological transformations working on the binarized input. Prior to the segmentation, we subject the inputs to an additional preprocessing and normalization step, which aims to enhance the contrast between motif and background regions.

Contrast Enhancement and Normalization. With a normalization step, we bring the raw object photographs into a standardized form so that the parameters of the subsequent segmentation can be tuned for a generalized input. *I.e.*, this involves scaling, removing noise, enhancing important, and discarding unimportant image information.

First, the images are scaled to a uniform height of 512 pixels which was established to be an appropriate trade-off between efficiency and preservation of low-level details. After converting the image to grayscale the primal-dual denoising algorithm by Chambolle and Pock [166] is applied to reduce noise while preserving essential image features. Subsequently, the contrast between motifs and background is maximized. Similar to the method by Otsu [167], we assume the presence of two dominant color classes in the image pixels belonging to the object surface. The assumption is based on a characteristic of ancient Greek pottery, whose surface paintings generally exhibit a bichrome colorization (SEC. 1.1). Hence, the two intensity peaks of the grayscale histogram of the masked object image (FIG. 3.4, top) should correspond to the mean color of the vessel surface μ_1 and the mean color of the motif paintings μ_2 , respectively. μ_1 and μ_2 are estimated with a Gaussian Mixture Model (GMM) with two components fitted to the histogram by maximizing the logarithmic probability $\ln P(H|\theta)$ of observing the histogram H given the parameters θ . Hence, the most likely model parameters $\theta_{ML} = \{\mu_1, \mu_2, \Sigma_1, \Sigma_2\}$, with covariances Σ are given by $\theta_{ML} = \arg\max_{\theta} \{\ln P(H|\theta)\}$. All values in-between μ_1 and μ_2 are spread linearly across the whole spectrum with values below or above that being clipped to the respective minimum and maximum intensities (FIG. 3.4, bottom). In the majority of cases, this process results in a significant contrast enhancement between motif and background pixels.

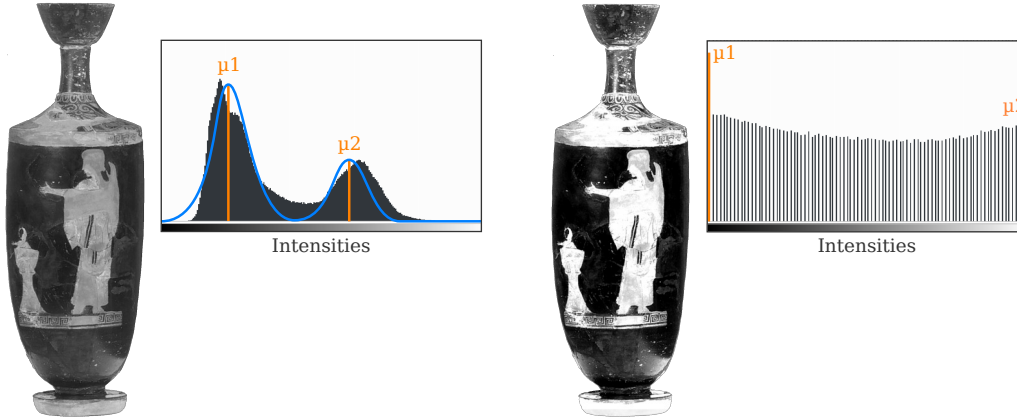


FIGURE 3.4. Left: the mean background and mean motif color can be observed as two intensity peaks in the logarithmic grayscale histogram. The locations of them (μ_1 and μ_2) are determined by fitting a statistical model to the histogram. Right: Spreading the histogram between μ_1 and μ_2 over the whole intensity spectrum greatly enhances the contrast between motif and background.

Graph-cut based Segmentation. While the graph-cut segmentation by Rother *et al.* [164], used for the object segmentation (SEC. 3.2.1), is appropriate for a binary separation of foreground and background, it is not able to perform an instance segmentation. Such can be achieved with the Efficient Graph-Based Image Segmentation (EGBIS) algorithm described by Felzenszwalb and Huttenlocher [165]. The behavior of this unsupervised segmentation can be governed by adjusting the minimum size of a component $|C|$ (in number of pixels) and the dimensionless scaling parameter k , which models a tendency towards larger components. We adjust these parameters empirically for our image data by fine-tuning them on a varied subset of images (FIG. 3.5, top row). A good segmentation for most inputs is obtained with $|C| = 1,500$ and $k = 1,500$ (FIG. 3.5, 2nd row). In order to preserve fine details, the resulting segments are further refined individually. To this end, a segment's mask is rescaled to the resolution of the original input image and used as a seed for the FG/BG graph-cut segmentation by Rother *et al.* [164], similar to the object segmentation. Finally, segments with an extent of less than 1% of the image size are classified as segmentation artifacts and discarded.

Morphological Segmentation. Although EGBIS generally performs well for our data, it is specifically sensitive *w.r.t.* small gaps, *i.e.*, lines in the gradient image resulting from the fine stylistic strokes exhibited by specific pottery art (FIG. 3.6(a)). As shown in FIG. 3.6(b), this can lead to an over-segmentation, separating the depicted Eros motif from the belonging wings. The dichromatic appearance of ancient Greek pottery makes it specifically suitable for a reduction to a binary representation. Thus, allowing for a straightforward extraction of connected foreground components. Directly labeling the binary image of the normalized grayscale output already shows a better preservation of the figure's shape but has similar problems in the face of gaps, producing a more complex silhouette as is apparent with the stylized feathers in FIG. 3.6(c). To this end,



FIGURE 3.5. A variety of vessels exhibiting motif paintings (top) together with the instance segmentation results obtained with EGBIS (middle) and morphological segmentation (bottom). With the latter, segments attributed to the surface background are marked with a black label.

we apply a series of morphological operations to the initial binary image. In the first step, we downscale the images to half the resolution for efficiency and apply a closing operation using a small disk-shaped structure element (2 pixel radius) to fill out fine recurring gaps in the motifs. As shown in FIG. 3.6(d), this reattaches the leg of the figure and closes the silhouette of the wings. In the second step, we add a morphological opening using a larger kernel (6 pixel radius). This has two effects: First, it produces a safer detachment of individual motifs. Second, it smoothes the silhouette and removes fine high-frequency details like the Kithara element held by Eros in FIG. 3.6(e). A connected component labeling [149] with 8-way connectivity is used to identify coherent regions. Similar to the EGBIS approach, small residual segments below a minimum threshold size (1% of the image area in all examples) are discarded.

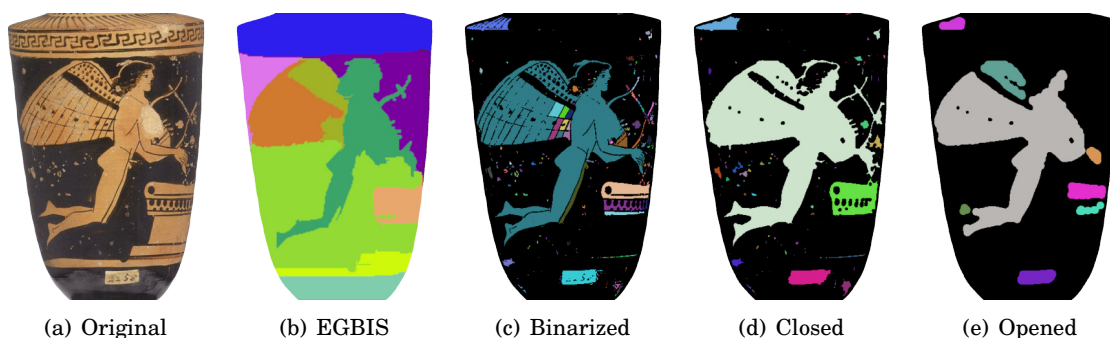


FIGURE 3.6. Comparison of different segmentation approaches for an Eros motif (a). The EGBIS is sensitive *w.r.t.* small gaps in the gradient – *i.e.*, the motif’s wings area. With an alternative approach, the input is binarized with a thresholding (c) before it is subjected to a morphological closing (d) in order to fill small gaps and a final morphological opening (e), removing unwanted noise. The colorization in (c)–(e) is the result of a connected component labeling.

Exemplary results for this segmentation technique are given in the bottom row of FIG. 3.5, which also shows the results obtained with the EGBIS. It can be seen that neither approach performs optimally on all of the inputs, but either one of them has its respective strength and weaknesses. Also, note that the segmentations do not distinguish between motif segments and complementary background segments. Yet, classifying image segments as regions of interest and non-interest is not trivial, and also not desirable, as it is without the scope of the segmentation process but the task of the applications building upon these segmentations.

Part I

Shape-based Retrieval

3D-BASED AND 3D+SKETCH-BASED RETRIEVAL

Contents

4.1	Introduction	42
4.1.1	Information Retrieval	43
4.1.2	Our Contribution	45
4.2	Related Work	46
4.3	3D to Normalized Profile Curve	53
4.3.1	Concept	53
4.3.2	Results	55
4.3.3	Discussion	57
4.4	Sketch-Aided 3D Retrieval	57
4.4.1	Concept	58
4.4.2	Query Preparation	58
4.4.3	Content-based Image Retrieval	62
4.4.4	Results	66
4.4.5	Discussion	77
4.5	Conclusion	79

The retrieval of objects based on shape was one of the core objectives of the CrossSAVE-CH project and remained thus an essential research focus throughout the project duration. Since the subject of shape-based retrieval, a subtopic of CBIR has been extensively researched for decades, our methodologic contributions to this subject are accompanied by an extensive related work section besides.

This chapter is based on:

- Sipiran, I., Lazo, P., Lopez, C., Jimenez, M., Bagewadi, N., Bustos, B., Dao, H., Gangisetty, S., Hanik, M., Ho-Thi, N.-P., Holenderski, M., Jarnikov, D., Labrada, A., Lengauer, S., Licandro, R., Nguyen, D.-H., Nguyen-Ho, T.-L., Perez Rey, L. A., Pham, B.-D., Pham, M.-K., “Shrec 2021: Retrieval of cultural heritage objects”, *Computers & Graphics*, vol. 100, pp. 1–20, 2021. DOI: [10.1016/j.cag.2021.07.010](https://doi.org/10.1016/j.cag.2021.07.010)
- Lengauer, S., Komar, A., Labrada, A., Karl, S., Trinkl, E., Preiner, R., Bustos, B., Schreck, T., “Sketch-Aided Retrieval of Incomplete 3D Cultural Heritage Objects”, in *Eurographics Workshop on 3D Object Retrieval*, The Eurographics Association, 2019, pp. 17–24. DOI: [10.2312/3dor.20191057](https://doi.org/10.2312/3dor.20191057)
- Lengauer, S., Komar, A., Labrada, A., Karl, S., Trinkl, E., Preiner, R., Bustos, B., Schreck, T., “A sketch-aided retrieval approach for incomplete 3d objects”, *Computers & Graphics*, vol. 87, pp. 111–122, 2020. DOI: [10.1016/j.cag.2020.02.001](https://doi.org/10.1016/j.cag.2020.02.001)
- Karl, S., Houska, P., Lengauer, S., Haring, J., Trinkl, E., Preiner, R., “Advances in digital pottery analysis”, *it - Information Technology*, 2022. DOI: [doi:10.1515/itit-2022-0006](https://doi.org/10.1515/itit-2022-0006)

4.1 Introduction

Shape is the primary property for the analysis of excavated pottery artifacts [168]. That is, archaeologists are expected to reconstruct the profile of a whole vessel from a single surviving fragment, allowing one to affiliate an artifact with a particular epoch, culture, and even manufacturer [64]. The classical way of achieving this is the pairwise visual comparison of an excavated object with already published artifacts. The objects described in related literature are commonly documented with a drawing of their profile curve, which is (apart from clay type and painting style) the basis for the comparison. As the number of pottery objects is enormous, the manual comparison of all objects is not feasible, but domain researchers have to preselect from a smaller collection of publications that are, according to their knowledge, appropriate for the object in question. This process has several downsides. It is very time-consuming and requires skilled personnel, as the domain researcher must maintain an encompassing knowledge base of related publications. Yet, at the same time, there is no guarantee that potentially relevant findings will slip through the radar, as it has already happened frequently. That is, oftentimes, fragments belonging to the same artifact are strewn across several museums or collections. This state, referred to by *Disiecta Membra* (eng. “scattered fragments”), results from artifacts being sold by grave robbers or other unconventional circumstances [10, 169]. Rediscovering a common affiliation of scattered fragments is a very challenging task.

In short, the need for a degree of automation in the similarity retrieval of pottery is evident. The state-of-the-art method of documentation is a high-resolution 3D model obtained with appropriate capturing hardware. In general, a push towards this kind of preservation could be

observed in the last two decades [168], with an ever-growing amount of digitized vases. Hence, existing 3D retrieval methods could be adopted to serve the archaeological use case. However, for the vast majority of objects, only textual, photographic, or drawn information is available. In order to include these objects in a similarity search, a retrieval system needs to be able to cope with the various different data modalities (*c.f.* SEC. 2.2) used in the CH domain. Tailored methods are necessary to bridge this modality gap and efficiently search through a large amount of data.

Another challenge characteristic of pottery research is the incompleteness of objects. Due to millennia of wear, very few excavated objects are whole, but the vast majority exhibit some degree of fragmentation or are just a pile of sherds in extreme cases. That is, retrieval methods also need to be able to cope with partial inputs, commonly known as *partial retrieval* in literature.

4.1.1 Information Retrieval

The process of Information Retrieval (IR) deals with the gathering of *relevant* entities from a database in an effort to address a *user information need*. To this end, this information need must be translated into a *query* – a representation understandable by the retrieval system – to be processed by a search engine. As a result, a subset of the database’s entities, which are determined to be relevant *w.r.t.* the given query, are presented to the user. The core difference to *data retrieval* is that (1) the obtained results are sorted by relevance, and (2) the results can be inaccurate as the query can be vague. In contrast to data retrieval, where one false positive in the results signifies a failure of the system, this is no reason for concern in an information retrieval system. On the other hand, having all entities satisfying the query (data retrieval) does not necessarily satisfy the user’s information need [170, 171].

In order to measure the similarity of an entity in the database to the query, we need a *logical view* of said entity. For text documents, which are by far the most common modality for IR, logical views are typically comprised of a set of *keywords* – words describing the document in a meaningful manner, either obtained automatically using text mining approaches or manually by a specialist. However, many modern archives do not only contain text documents but can comprise a series of other data modalities, such as images (gray or colored), videos, graphs, sound (digitized voice or music), time series (stocks or scientific measurements) and many more (*c.f.* SEC. 2.2). The IR performed on such multi-modal archives is referred to as *Multimedia IR*. The queries used in Multimedia IR can be *attribute-based*, focusing on an item’s associated metadata or labels (*e.g.*, “give me photographs which were taken in 2022”), or *content-based* – *i.e.*, Query by Example (QBE) [170, 171]. Prior is unfortunately not feasible for large-scale multimedia repositories as the necessary annotated information is problematic to extract automatically but has to be generated manually. In the context of our work, we focus exclusively on the latter, the problem of Content-based Multimedia IR, where multimedia data itself is used as a query [87]. Within this class of IR, two (extensively researched) subcategories are of particular interest: Content-based Image Retrieval (CBIR), querying a database of images by example, and 3D Object Retrieval

(3DOR), querying a database of 3D objects by example.

Content-based Image Retrieval. The topic of CBIR has been extensively researched for the last two decades. The basic idea is to match the visual content of a query image to the images in a database. To this end, images need to be transformed such that they are represented by a vector of numerical values in a meaningful fashion. The process of obtaining such a Feature Vector (FV) is the objective of a Feature Descriptor (FD), which generally aims to reduce the semantic gap between an image's FV representation and a human's visual understanding. As there is no universal definition of what makes images similar or dissimilar from a human perspective, there are infinitely many possibilities for designing a FD. Very many concepts have been published to date. According to Latif *et al.* [172], many of them use low-level visual features, such as color, shape, texture, and spatial layout, or combinations thereof. Additionally, we can differentiate between global features, which are computed from the whole image, and local features, which rely on 'significant' points in the image differing noticeably from their neighborhood. Most modern CBIR approaches rely on deep learning techniques where CNNs, in particular, have demonstrated outstanding capabilities of capturing the essence of an image.

3D Object Retrieval. Researchers to date have investigated many approaches for 3DOR based on various similarity concepts such as shape, structure, appearance, or metadata [173, 174]. Yet, not all approaches are suitable for any retrieval task. Quite the contrary, the selection of an appropriate method depends on several factors, such as (i) efficiency (if the queried repositories are large); (ii) tolerance *w.r.t.* noisy or fragmented data; (iii) completeness of data (*c.f.* partial retrieval); and also (iv) the representation of 3D objects (there is no uniform representation for a 3D object but several options with respective advantages and disadvantages SEC. 2.2.5). For most applications FDs should be invariant *w.r.t.* differences in orientation, scale, level of detail, and location of a 3D model since the embedding in 3D space is usually no desired notion of similarity. That is, the FVs obtained from similar shapes should also be similar, independent of said factors. Some FDs achieve this requirement by design if they build upon relative object properties, *e.g.*, distribution of surface curvature, while others require a preprocessing step to transform all 3D objects into a normalized representation. This normalization typically starts with an alignment with the reference coordinate system using a Principal Component Analysis (PCA), a translation of the object's center of mass into the coordinate system origin, a canonical scaling (*e.g.*, to fit into the unit cube), and (optionally) a flipping based on a moments criterion [87]. Based on the FD type, the normalized or original object is abstracted by a volumetric, surface, or image representation. For the image abstraction, one or more images from specific views are generated using rendering techniques and representing either object silhouettes, depth maps, or the textured object. According to Bustos *et al.* [87] FDs can be categorized into statistics (basic object properties), volume-based, surface geometry, or image-based.

The recent success of deep learning models for image classification, and retrieval in general, has prompted for application of such models also for 3DOR. Two main approaches are followed to input 3D object data to deep networks, representing 3D models either as a set of 2D views or as a voxel grid. Generally, models that use 2D views surpass voxel-based models. However, the latter may perform better if more complex neural network models are used, being much more expensive to train and also needing a substantial amount of relevant data for training. In particular, trained deep models have been used for feature extraction, yielding *learned* features [175]. For in-depth descriptions of techniques, we refer to the surveys of methods using deep learning techniques on 3D models by Ioannidou *et al.* [176] and Gezawa *et al.* [146]. Recent results show that in the presence of sufficient training data, learned features can outperform engineered features in terms of retrieval and classification accuracy. However, in the absence of necessary training data, relying on engineered features incorporating domain knowledge about the expected target set is more feasible.

The query for a 3DOR system does not necessarily need to be a 3D object. Besides searching with a given 3D object within a set of 3D objects, one can also search with and within views (images) of 3D objects, drawings, sketches, or videos. Sketches in particular have the advantage that no query object or view is required and provide a natural user interface [69, 70]. Sketching also is connected to modeling, and previous works have proposed modeling based on example objects found by sketching or applying generative modeling methods detected from sketches [71]. If the query modality differs from the modality of the search space, they have to be transferred into a common base modality. View-based approaches are frequently used, as views can be a common denominator among different modalities.

Partial 3DOR. A subdiscipline of 3DOR, which is of particular interest for CH applications, is Partial 3DOR (P3DOR). It refers to the general tasks of ‘whole-to-part matching’ or ‘part-to-part matching’, with naïve solutions to both tasks having a $\mathcal{O}(n^3)$ complexity [38]. Partial retrieval is especially relevant for CH since excavated artifacts are usually fragmented. Pottery, in particular, is frequently preserved only as a set of sherds. Several methods for P3DOR have been proposed within the last years, as indicated by the survey of Savelonas *et al.* [177]. The general idea is to employ local shape information, which can be combined into a Bag-of-Visual-Words (BoVW) to derive global shape representations of local shape descriptors. Variants of this concept have been successfully used in diverse applications [67, 178, 179]. Nonetheless, P3DOR is considered a more complex problem than conventional 3DOR, emphasized by a benchmark by Sipiran *et al.* [180], who concluded that the problem is far from being solved.

4.1.2 Our Contribution

Regarding content-based retrieval, we present two orthogonal approaches for retrieving rotationally symmetric pottery artifacts based on different input modalities and underlying concepts. The

first concept relies solely on the 3D geometry from which a profile curve is extracted (SEC. 4.3). Said is used for the pairwise similarity comparison to other 3D models. The second approach (SEC. 4.4) is particularly well suited for retrieving incomplete 3D inputs. That is, the missing parts of such a fragmented input are completed with a user sketch and compared against a database comprised of photographs, leveraging geometric as well as textual information. Before that, a detailed investigation of related work on the subject of content-based pottery retrieval, together with a delineation *w.r.t.* to our work, is given in SEC. 4.2. Finally, SEC. 4.5 concludes the chapter by summarizing our contributions and the still open research questions.

4.2 Related Work

Over the decades, many methods specifically for the content-based similarity retrieval of ancient pottery have been presented. While textual searches over structured constitute the most traditional variant of retrieval, they are somewhat limited in terms of their applicability. Consequently, content-based approaches have received steadily increasing attention. Most of them leverage the rotational symmetry of pottery artifacts exhibited by objects across cultural and epochal boundaries due to their unique way of manufacturing (potter’s wheel). Due to this fact, most early works focus on comparing vessel profiles [20, 32–42, 59], which unambiguously describe a solid of revolution. One significant advantage of vessel profiles is the availability of data, as it is one of the oldest archaeological methods to capture a pottery object’s shape. Hence, we find an abundance of them across relevant publications to date. Their compact form for describing a shape makes them the perfect basis for several taxonomies of vessel types [181–184]. Later methods [5, 72, 73, 88–93, 185] are based on image data (primarily photographs) which carry additional textual cues like color, painting style, etc. Most recently, due to more readily available digitization hardware, 3D models, which comprise all of an object’s shape and texture information, have become the preferred data basis [60–67]. Also, cross-modal retrieval methods, where the query is comprised of a modal hybrid [72, 73] or where the modality of the search space differs from the modality of the search space [5, 20, 33, 36, 37, 72, 73, 88–93], have been investigated.

Not only can we group methods by the data modality (profile curves, photograph, 3D models) used for similarity computation but also by the type of query or they or the type of data the search space is comprised of. Regarding queries, the same data modalities, such as vessel profiles, photographs, videos, or 3D models, and more exotic examples like sketches or hybrids of multiple types have been investigated. With the rapid advancement of ML-based techniques (CNNs in particular) in all fields of computer vision, we can also distinguish between approaches using of deep learning approaches [5, 89–93] as opposed to ‘conventional’ feature-based concepts. Several approaches deal specifically with the retrieval based on an incomplete query object (P3DOR) [5, 32, 35, 38, 39, 41, 63–67, 89–93, 185]. TAB. 4.1 gives a structured breakdown of methods which are covered in more depth in the following.

TABLE 4.1. The taxonomy of 3DOR methods for CH applications, clustered by the data basis the algorithms are based on (vessel profile, image, 3D model). We also state the approaches' input modality in terms of query and search space, which comprise the modalities *Profile Drawing* (PD), *Photograph* (I), *Video* (V) and *3D Model* (3D). Our proposed concepts are marked in bold font.

Algorithmic Basis	Concept	Query	Search Space	Deep Learning	Preservation State [†] (Query Search Space)
Profile	Semantic knowledge graphs [32]	3D	3D		(B B)
	Rule-based matching [33]	3D	1-*		(C C)
	Non-rigid deformation analysis [34, 59]	PD	PD		(C C)
	Profile fitting with genetic algorithm [35]	3D	3D		(F C)
	Hierarchic PC features [20, 36, 37]	S	3D		(C C)
	Profile correlation [38]	PD	PD		(B B)
	Local shape features [39]	PD	PD		(B C)
	Discrete Silhouette [40]	I	I		(C C)
	Shape Harmonics [41]	PD	PD		(F C)
	Normalized Profile Curve (part of [42])	3D	3D		(C C)
Image	Shape and color matching [185]	I	I		(B* C)
	Local visual features [88]	3D	I		(C C)
	CNN Classifier [89]	V	I	✓	(B C)
	CNN Classifier [5, 90–93]	I	PD	✓	(F C)
	Sketch-completed Query [72, 73]	3D+S	I		(B C)
3D	Morphological features [60, 61]	3D	3D		(C C)
	Depth maps [62]	3D	3D		(C C)
	Panoramic projection [63]	3D	3D		(F C)
	BoVW with Local shape features [64–66]	3D	3D		(F C)
	Fisher encoding of Local shape features [67]	3D	3D		(F C)
	Geometric and photometric features [186]	3D	3D		(C C)
Benchmarks [68]		3D	3D	✓ ⁺	(F C)
Benchmarks [42, 187]		3D	3D	✓ ⁺	(C C)

[†] The preservation state [59] including complete or near complete (C), fragment (F) and both (B).

^{*} Geometric properties of vessel types specified in a rulebook [181].

^{*} Fragmented queries are possible if all fragments of a vessel are available.

⁺ A collection of methods is presented, some of which are relying on deep learning.

Profile-based. One of the earliest efforts towards shape-based similarity retrieval for (fractured) solids of revolution is presented by Sablatnig and Menard [32], who combine shape features with categorical metadata such as color, clay, or surface type for search and reconstruction purposes. To this end, they manually obtain the profile curve of a vessel fragment's 3D model, which is then matched against a database of archived profiles using a semantic network. In such a network, which is a labeled and directed graph, the nodes represent objects, sub-objects, or shape primitives, while the edges describe relationships between them. To determine which part of a profile curve belongs to a semantically cohesive object part, such as edge, border, or ringbase, they require input from an expert user. In a later work, Kampel and Sablatnig [33] show how they can automate this manual segmentation by detecting characteristic points along the profile, such as orifice points, inflection points, corner points, *etc.* [188]. From these points, descriptive shape properties of a fragment can be computed, which are compared against an existing taxonomy [181], stating geometric rules for individual vessel shapes of ancient Roman pottery.

Within the CATA (Archaeological Wheel Pottery of Andalusia) project¹, Martínez-Carrillo *et al.* [59] present an encompassing concept for the storage, analysis, retrieval, and visualization of archaeological pottery. To this end, they conceptualize a data model comprising structured metadata, such as information pertaining to the distribution, functionality, and state of preservation of a pottery object, together with different types of digital data, such as numerical values, text, images, and videos. Users can introduce their data into the curated database via a web interface. This interface allows uploading data and querying the database in a text or content-based manner. For the latter, a user can provide the vessel profile of an object he wants to search for. In terms of similarity computation of vessel profiles, the authors rely on a non-rigid deformation analysis which they presented in an earlier publication [34].

In a similar fashion, the Classical Art Research Online Services (CLAROS) project², a research initiative by the university of Oxford, aims to unite more than two million object records from research centers in Oxford, Paris, Cologne, and Berlin. The concept is inspired by the well-established Beazley archive [8] (*c.f.* SEC. 2.1.1) and similarly constitutes a database of archaeological photographs. However, besides digital archiving of Greek pottery, they also support a query by example retrieval with an image as a query. In terms of the underlying methodology described by Kurtz *et al.* [40], they proceed by automatically extracting the objects' silhouettes from canonical archaeological photographs using a FG/BG segmentation coupled with a symmetry detection to determine the mid-line. A silhouette is sampled in the vertical direction to obtain an FV containing the distances from the mid-line. To allow for efficient retrieval in the face of a huge database, they approximate the distances between a query and a database item and employ a random forest of k -D trees.

Within the Systeme d'Imagerie et d'Analyse du Mobilier Archeologique (SIMA) project, Maiza

¹<http://cata.cica.es/>

²<http://www.clarosweb.org>

and Gaildrat [35] study the comparison of vessel sherds (given as 3D models) to a database of whole 3D models. To this end, they leverage profile curves. For the whole models in the database those can be easily obtained by introducing a number of vertical cut planes since the models' rotation axes are already aligned with one of the coordinate axes. The profile curve of a sherd (query) is extracted in a similar fashion by using several cut planes after a manual alignment. From the resulting profile curves, only the longest one is used. While this input is sampled, given a 2D point vector, the whole models are described using implicit functions. The similarity computation between the input and a database element relies on an optimal alignment of profiles using a fitness function. That is, this function measures the distance of the sampled points to the implicit surface and allows thus to evaluate an alignment population's quality.

The 3D Knowledge and Distributed Intelligence (3DK) project³ is a collection of three sub-projects, focusing on (1) 3D morphology of ceramic vessels; (2) lithic refitting, the study of the development of stone tool technology; and (3) 3D topology of joint surfaces, which investigates the ability of human ancestors to make tools and walk upright [36]. Of particular interest *w.r.t.* 3DOR is the first topic dealing with the feature-based similarity retrieval of Native American pottery. That is, Razdan *et al.* [20] present an approach where they use hierarchic features obtained from the rotationally symmetric vessels' profile curves. From the dataset comprised of 3D models, these profile curves are extracted using curving planes. Cubic NURB curves are fitted to the intersection points. On those, they determine the four feature points of profile curves defined by Birkhoff [189] (*endpoints, points of vertical tangency, inflection points, and corner points*), which serve as the basis for a formal feature descriptor. Querying of the database is enabled through a visual interface which allows the user to provide a profile sketch alongside traditional text search.

Other recent profile-based approaches, specifically for fragmented inputs, include the work by Hristov and Agre [38]. They extract the inner and out profiles of archaeological drawings using a series of simple heuristics coupled with basic image processing techniques (binarization, symmetry detection, *etc.*). They also support the retrieval based on fragmented inputs, which are anchored to the bottom or top before a correlation coefficient *w.r.t.* to the other entries in the database is computed. Piccoli *et al.* [39] leverage the geometric properties of profile curves together with visual features (color and texture) to match a newly obtained sherd with the entries in a reference catalog. To this end, they use two established taxonomies of profiles for the search space and evaluate their approach based on a ground truth example compiled by experts. Their matching algorithm is inspired by the SIFT [95] approach to ensure invariance *w.r.t.* translation, scale, and rotation. But instead of gradient information of an image, they capture the 2D shape of the profiles with keypoints. Wilczek *et al.* [41] present a new way of capturing the essence of a profile – the Discrete Cosine Transform. This Fourier-based method transforms open outline coordinates into a set of harmonics, each composed of a pair of coefficients that may be used

³<http://3dk.asu.edu/>

as shape descriptions. They compare their method to three established existing approaches using synthetic queries, which were produced by synthetically fracturing vessel profiles given in a paper catalog. Based on this evaluation, the authors conclude that all of the compared algorithms perform very well but differ in flexibility and computation time and that the choice of an algorithm thus heavily depends on the intended application.

Besides retrieval, profile curves are also used to introduce a shape-based clustering to a collection of objects. Such can be achieved using a mathematical representation of the profile describing the radius, tangent, and curvature along the arc, as proposed by Karasik and Smilansky [190]. Regarding deep learning, Navarro *et al.* [191] show how ceramics can be automatically classified using transfer learning for retraining a Residual Neural Network (ResNet). At the same time, Parisotto *et al.* [192] leverage non-linear learned features from the latent space of a deep convolutional variational autoencoder network.

Image-based. Bishop *et al.* [185] are among the first to combine shape with textual information to retrieve ancient Greek pottery. Archaeological photographs serve as a basis for the search space, which can be queried, also using photographs. For the retrieval based on whole vessels, the object masks of the depicted object are obtained using a FG/BG segmentation. From the mask, regional properties, *e.g.*, the eccentricity of an enclosing ellipse, are combined with color properties to match the query to the search space entities. A partial search with a sherd as the query is also supported to some extent. Moreover, if all parts of a shattered object are available, the individual photographs are put together using a jigsaw puzzle algorithm before the same retrieval as with the unbroken inputs is conducted.

Albeit not using it for retrieval purposes but for unsupervised classification, Johácz [13] shows how geometric morphometrics can be used for describing the shape similarity of vases. To this end, she uses both an outline analysis of object masks as well a set of 15 manually defined landmarks on the 3D shape. The evaluation on a set of 30 Attic lekythoi – a vase type known to be well described by its shape – revealed that the prior method is better suited for practical applications and results in a proper clustering.

Gregor *et al.* [88] introduce a preliminary approach for searching views of 3D CH objects in image data using the local MPEG-7 visual descriptor along with other established local descriptors.

Tyukin *et al.* [89] outline a concept that allows turning commodity video-capturing devices, such as smartphones or tablets, into classifiers for whole vessels as well as sherds. To this end, they use a CNN, fed with HoG features extracted from the video feed to cope with the limited given hardware. A downside of this approach is the potentially vast number of training samples with samples from different angles and lighting conditions for a robust classifier. Nevertheless, the authors show that they are able to achieve meaningful results with a minimal initial dataset.

Núñez Jareño *et al.* [90] similarly use deep learning techniques to classify Roman pottery. However, they address the issue of sparsely available training data by employing a transfer

learning concept. An initial dataset of several thousand labeled photographs of complete and near-complete vessels is complemented with synthetically generated images. For the latter, solids of revolution are generated from a taxonomy of profile curves which are subjected to different degrees of synthetic erosion to mimic the traits of wear. The authors show that with their approach, generating a single image classifier with acceptable accuracy is possible.

The ArchAIDE project⁴ [5, 91–93] follows along the same line. The ultimate objective is to develop a mobile tool that supports archaeologists with an on-site classifier for pottery sherds based on a single image. They use a machine learning classifier that trains on synthetically produced training data. To this end, they create a set of 3D models from profile curves extracted from archaeological drawings. To produce sherds from these models, an off-the-shelf Voronoi fracturing approach is used.

3D-based. One of the first approaches leveraging information from a 3D model for classifying ancient pottery is presented by Hörr *et al.* [61], who define a hierarchy of features to this end. That is, primary features, for example, are global object features such as a vessel’s absolute height, diameter, *etc.* Even though they are trivial to extract, they are oftentimes sufficient for determining the basic shape class (amphora, tureen, jug, pot, *etc.*). For the extraction of secondary and tertiary features, the authors propose a segmentation pipeline that, first of all, removes all appendages before the main vessel parts (base, belly, shoulder and neck) are determined using the profile curve.

Koutsoudis *et al.* [60] present a novel shape descriptor for complete or nearly complete 3D vessels based on morphological features. One of the main contribution of their work is the development of a tailored pose estimation algorithm that is able to detect a solid of revolution’s rotation axis even in the presence of appendages such as handles, feet, or lugs. This pose estimation is applied in the course of the input normalization step, which scales an input to the unit sphere and aligns its rotation axis with the *Y*-axis. For extracting the compact FV, properties like the radius and center coordinate of a circle fitted to the mesh contours at different *Y*-offsets are determined. They show that the proposed descriptor substantially outperforms the established MPEG-7 descriptor for retrieving ancient pottery vessels. A year later, Koutsoudis and Chamzas [62] showed how depth maps could be used instead of the morphological features to the same end. They perform the same normalization steps and extract two orthogonal depth maps encoded into a compact FV with two different established FDs.

Sfikas *et al.* [63] present a method for retrieving a partial 3D CH object based on local features. The database of complete 3D objects is subjected to a panoramic depth projection from which Dense SIFT features are computed. Similarly, a range of image representations is computed in order to bridge the representation gap. With a set of qualitative examples, the authors show that the retrieval quality is not severely affected by missing 3D geometry.

⁴<http://www.archaide.eu>

Within the Tepalcatl project, Roman-Rangel *et al.* [64] investigate the automatic categorization of ancient Mexican potsherds. To this end, they employ two state-of-the-art local 3D descriptors (Scale Invariant Spin Images [193] and Local Depth SIFT (LD-SIFT) [193]), which they implement using a BoVW approach. Based on their experiments, they conclude that LD-SIFT performs better in all observed cases. In a later publication [65, 66], they develop a new local 3D shape descriptor on their own, referred to by *histogram of spherical orientations*, which they use again in combination with a BoVW approach. The idea of the descriptor is to capture the distribution of spherical orientations of a vertice’s neighboring points. The authors claim that this descriptor is scale-invariant and highly robust *w.r.t.* rotation and noise and prove its effectiveness by comparing it with Spin Images and SIFT as well as seven other established feature descriptors.

Savelonas *et al.* [67] employ a Fisher encoding [194] of local 3D features for the P3DOR on the Hampson pottery dataset [120]. Their proposed method uses a hybrid shape-matching scheme, incorporating both local and global shape similarities for multiple scales. Even though it achieves state-of-the-art retrieval performance, the authors state that the results remain far from perfect and that the task of P3DOR for CH is still an open problem.

Biasotti *et al.* [186] consider both geometric properties, such as curvature, size, roundness, and mass distribution, as well as photometric properties, such as texture, color distribution, and reflectance, for measuring object similarity. To this end, they investigate different basis functions for describing the shape of a 3D model and different colorimetric features, which are given as real and vector-valued functions. Similarly to Savelonas *et al.* [67], they use a set of objects from the Virtual Hampson Museum to evaluate the effectiveness of their proposed measures for both retrieval and classification purposes.

Benchmarks. Three noteworthy benchmarks for the objective of 3DOR for CH applications have also been presented. “The Orion Pottery Repository” by Stergioulas *et al.* [187] comprises 160 textured 3D digital replicas of ancient Greek pottery, which are also classified using an in-house developed thesaurus.

The other two examples are entries to the long-standing Shape Retrieval Contest (SHREC)⁵, a series of annual contests for evaluating methods and algorithms related to 3D object retrieval. First, the 2016 SHREC track on “Partial Shape Queries for 3D Object Retrieval” by Pratikakis *et al.* [68], who provide a P3DOR challenge with objects of varying scan qualities and degrees of partiality. The selected dataset consists of 383 models from the Virtual Hampson Museum collection [120]. Four research groups contributed to the challenge with eight distinct algorithms. The queries are both artificially generated and obtained from real objects. For the latter, scans of varying quality are used, which proved to be very challenging for all submitted approaches. In the 2021 SHREC track on “Retrieval of cultural heritage objects” by Sipiran *et al.* [42], contestants were asked to conduct a retrieval on a given set of complete or near complete pottery artifacts

⁵<https://www.shrec.net>

from pre-Columbian cultures in Peru, exhibiting multifaceted shapes and varied artistic styles. To this end, a training dataset with about 1,000 objects from the same corpus was provided. The majority of the 30 submitted approaches were learning-based, which shows a clear predominance over engineered methods in the face of sufficient training data. Nonetheless the best approach scores a mean average precision value of 0.8604, indicating that the problem is far from being solved.

Positioning of our work. As visible in TAB. 4.1, our proposed methods (Normalized Profile Curve and Sketch-completed Query) fill spots of still unprecedented combinations of technique as well as query and search space modality. Using profile curves for retrieving pottery is not unprecedented, but we present a novel and robust method for extracting said curve from a (potentially fractured) 3D input.

Regarding the second method, the novel and unique query modality – a hybrid of a 3D model and user sketch – is noteworthy. While neither P3DOR nor sketch-based retrieval is unheard of, we believe that the combination of the two provides an intuitive way to incorporate user knowledge in the query formulation while still using all available original object information. Moreover, this query is presented in the context of a cross-modal retrieval approach – 3D plus sketch as query and images in the search space. The image modality of the search space is chosen deliberately to address the reality of data availability. The most influential preparatory work for this concept is the idea proposed by Gregor *et al.* [88]. While building on this work, our method additionally (i) addresses the problem of incomplete queries; (ii) evaluates the concept much more rigorously with a significantly larger search space and automated quantitative tests; and (iii) uses a more carefully selected set of FDs based on extensive experiments.

4.3 3D to Normalized Profile Curve

With the Normalized Profile Curve (NPC) descriptor, we leverage the rotational symmetry properties of vessels in a similar fashion to the majority of retrieval approaches for pottery. The concept is loosely based on Kurtz *et al.*'s [40] profile curve descriptor (*c.f.* SEC. 4.2) but adapted for 3D inputs. It was submitted to the 2021 SHREC track on “Retrieval of cultural heritage objects” [42] (*c.f.* SEC. 4.2), where its effectiveness was evaluated on a dataset comprised of pre-Columbian pottery.

4.3.1 Concept

While there have been several approaches that extract profile curves from a 3D model [32, 33, 35], they usually rely on mesh contouring of the properly aligned mesh in order to obtain a profile curve. They either use the contour yielding the longest contour line [33] or extract several contours based on different angular cuts [35]. This approach would also be possible with the

given dataset since it is comprised of textured coarse triangle meshes, but we found that, due to both the high degree incompleteness of some inputs as well as a variety of appendages attached to other inputs, such contours are often erroneous and do not capture the profile curve faithfully.

We propose a different concept. Disregarding all topology information, we take the mesh vertices as the sole input. The input should be aligned so that its center of mass coincides with the origin of the 3D space and its rotation axis coaligns with one of the Cartesian coordinate axes. However, this normalization was not necessary for the given dataset as the given models were already preprocessed this way⁶ (the rotation axes are aligned with the Y -axis). Yet, some models' variations in scale, as well as fragments of inner vessel surfaces in some models, necessitate further normalization efforts. Hence, all objects have been subjected to a series of preprocessing steps comprised of: (i) removing unreferenced vertices, (ii) fitting the model to a unit cube, (iii) shifting for 0.5 in Y -direction, and (iv) removing vertices belonging to the inner vessel surface (vertices whose normals point towards the rotation axis). For each vertex $v = \langle v_x, v_y, v_z \rangle$, we compute the distance to the rotation axis $r_v = \sqrt{v_x^2 + v_z^2}$, which is comparable to a projection in the angular direction. These distances are sorted into B ($B = 50$ used for the submitted results) equal-sized bins along the rotation axis and used for determining the bin-wise mean radii vector $\bar{\mathbf{r}} = (\bar{r}_0, \dots, \bar{r}_{B-1})$ (mean distance from the rotational axis) and the corresponding variances $\sigma = (\sigma_0, \dots, \sigma_{B-1})$. Note that handles, holes, or other surface irregularities of the model do not substantially influence the mean profile curve (FIG. 4.1) since the majority of projected vertices are in close vicinity to the profile curve. The similarity of the NPCs of two inputs is given by the Euclidean distance (EQ. (4.1)).

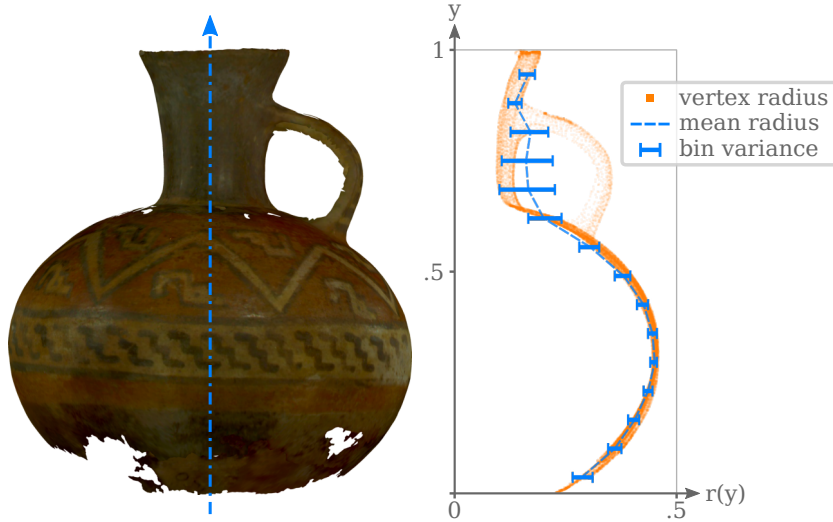


FIGURE 4.1. The vertices of the preprocessed mesh of a jar object (left) are projected along the angular direction (right). The profile curve is described by the binned mean radii $\bar{\mathbf{r}}$ and variances σ .

We achieve a performance improvement by assigning static weights $\mathbf{w} = (w_0, \dots, w_{B-1})$ to the

⁶<http://www.ivan-sipiran.com/shrec2021.html>

mean radii with $\mathbf{w} \propto \bar{\sigma}$, based on the assumption that bins with low variance across all objects are more discriminative than others. $\bar{\sigma} = (\bar{\sigma}_0, \dots, \bar{\sigma}_{B-1})$ denotes the bin-wise mean-variance with $\bar{\sigma}_b = \frac{1}{N} \sum_{i=0}^N \mathcal{S}[i]_b$ as the mean-variance for the axial bin b and $\Sigma = \{\sigma_0, \dots, \sigma_{N-1}\}$ as the set of all profile curve variances of a search space with N objects. We refer to this descriptor variant as Weighted NPC (W-NPC).

4.3.2 Results

The SHREC track consists of two separate challenges, ‘Retrieval by Shape’ and ‘Retrieval by Culture’. In those, the W-NPC descriptor competes against 9 and 7 other (deep learning-based) approaches, respectively. Various standard performance measures, including Nearest Neighbor (NN), First-tier (FT) precision, Second-tier (ST) precision, Mean Average Precision (mAP), and Normalized Discounted Cumulative Gain (nDCG), are used for comparison. `tabreftab:sketch-plus-3d:evaluation-metrics` presents the values our approaches scored on the test set.

TABLE 4.2. NN, FT precision, ST precision, mAP, and nDCG of the NPC and W-NPC descriptor, scored on the ‘Retrieval by Shape’ and ‘Retrieval by Culture’ challenge (*Values source: [42]*).

Challenge	Method	NN	FT	ST	mAP	nDCG
Retrieval by Shape	NPC	0.6751	0.3766	0.3106	0.4019	0.7874
	W-NPC	0.6787	0.4151	0.3237	0.4274	0.7996
Retrieval by Culture	NPC	0.7037	0.5489	0.7539	0.5736	0.8267
	W-NPC	0.6772	0.5523	0.7565	0.5776	0.8273

In the case of the Retrieval by Shape challenge, the W-NPC clearly outperforms the NPC achieving higher values in every evaluation metric. FIG. 4.2, left, shows the precision-recall plots based on the best run (highest mAP) of all submitted methods. The confusion matrix (FIG. 4.3, left) shows no abnormality besides the expected peaks on the main diagonal.

With the Retrieval by Culture, no dominance of the W-NPC over the NPC is observable, which we attribute to the fact that this is a generally much more complex challenge for which our proposed descriptor is far from optimal. The still overall higher values for this challenge result from the highly asymmetric distribution of classes. While the dataset for the Retrieval by Shape challenge features between 34 and 221 objects per class, the Retrieval by Culture class has one class (Culture2) containing 455 objects and 5 other classes with only 7 to 61 objects. This asymmetry can also be observed in the respective confusion matrix (FIG. 4.3, right), which shows that many objects are erroneously classified as not-belonging to Culture2. Moreover, we attribute the weaker performance of our descriptor for this challenge to the fact that the assumption of rotational symmetry does not hold in many cases since the dataset for this challenge is more faceted and also contains several other shapes like figurines. The authors of the study hypothesize that an object’s belonging to a culture is reflected in both its geometry and color. However, preliminary experiments where we combined our shape descriptor with color features

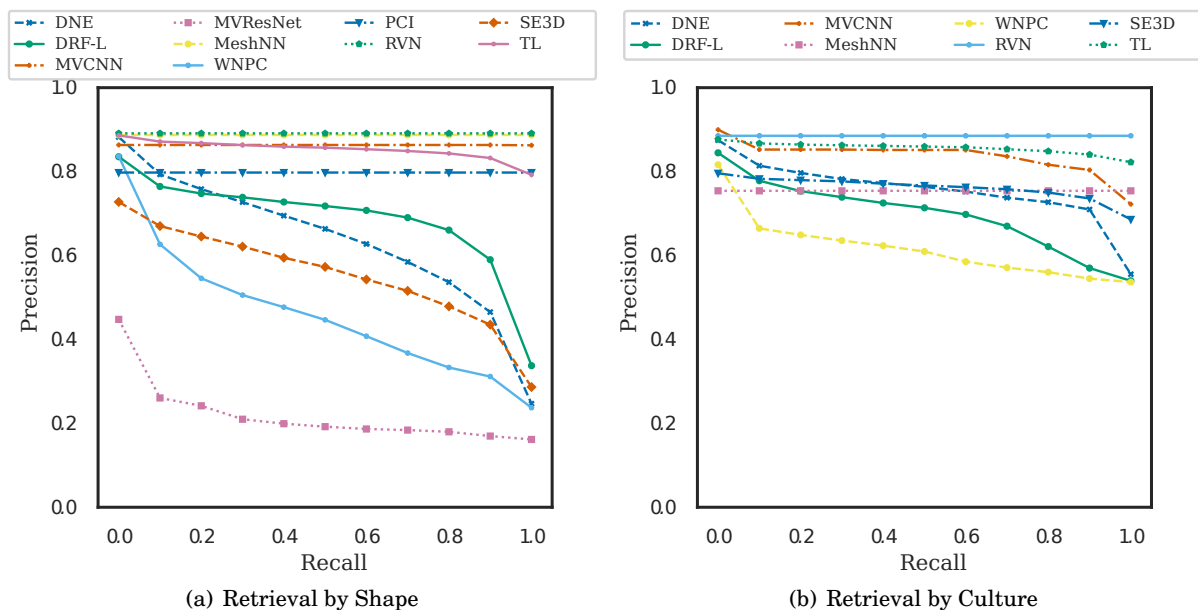


FIGURE 4.2. Precision-recall plots for the best run (determined by mAP) of every submitted method for the (a) Retrieval by Shape and (b) Retrieval by Culture challenge. (Image source: [42])

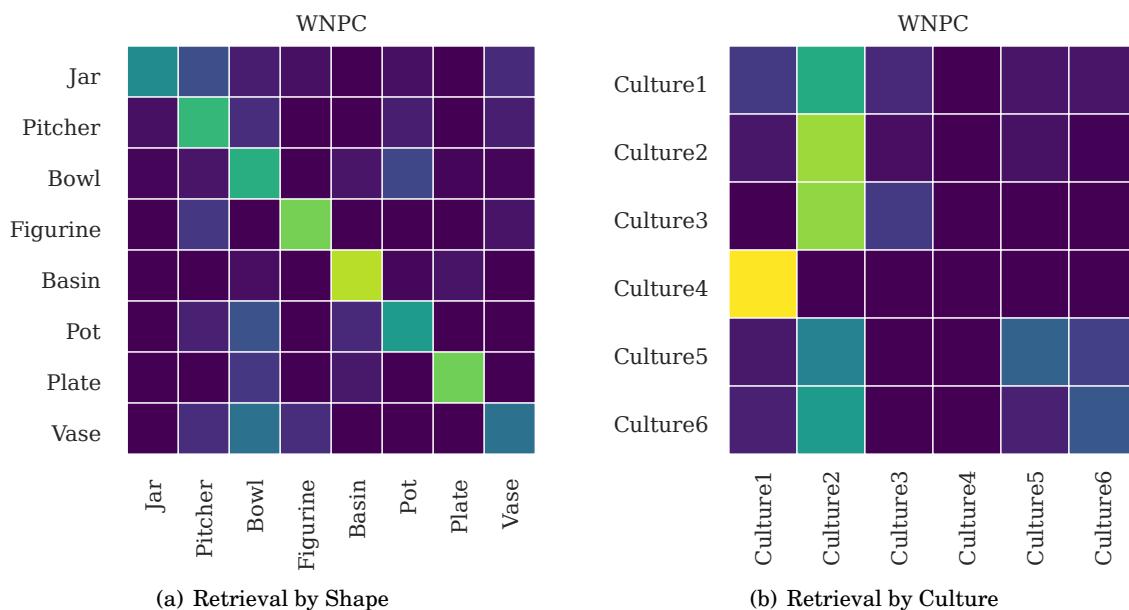


FIGURE 4.3. Confusion matrices for the (a) Retrieval by Shape and the (b) Retrieval by Culture challenge of the best run (determined by mAP) of the W-NPC. (Image source: [42])

did not improve performance. Also, two of the three best-fairing methods rely solely on silhouette information. FIG. 4.2, right, shows the precision-recall plots for this challenge.

Regarding efficiency, methods are hard to compare since every group uses different setups

and degrees of optimization. Nonetheless, we want to state the computation time we observed on commodity hardware. The offline processing involving basic geometry processing tasks is conducted mainly with MeshLab⁷ and takes about an hour for the whole dataset. A naïve implementation for calculating the profile curve and variances takes 11.9ms per object on average while computing the L_2 distance between two objects took 14.33 μ s on average.

4.3.3 Discussion

Even though most other methods clearly outperformed the proposed descriptor (at least in the Retrieval by Culture challenge), we believe it still holds a valuable contribution. That is, the NPC/W-NPC is the only submitted method that does not rely on deep learning techniques, making it comparatively easy to employ and transparent. Also, we want to note that the descriptor is not specifically designed or trained for the given dataset, making it readily applicable for retrieving any 3D shapes exhibiting rotational symmetry.

4.4 Sketch-Aided 3D Retrieval

We propose a new query modality, a hybrid of 3D views and user sketches (3D+Sketch). This allows to provide additional context to a query, which is particularly useful for searching based on incomplete query objects (partial retrieval) or for testing hypotheses regarding the existence of certain shapes in a collection. Compared to purely sketch-based approaches, our concept has the significant advantage that all the preserved texture and shape information is used to support the retrieval process. To this end, we present an appropriately designed workflow for constructing meaningful queries from incomplete 3D objects together with a user sketch suggesting missing parts. Additional visual cues help users compare retrieved objects with the query. The closest predecessor to this work is the retrieval approach by Gregor *et al.* [88], who also use 3D models for searching in a database of images. While building on the same problem, our work presents a much more encompassing search approach considering geometry, texture, complementary user sketches, and useful interface extensions for query weighting and result visualization. With region-based query weighting, we allow users to interfere with the similarity computation since archaeological research is often interested in finding object analogies, emphasizing or de-emphasizing certain characteristics. Using both a qualitative evaluation with a few characteristic use case scenarios as well as a quantitative evaluation based on synthetically generated queries, we demonstrate the feasibility and potential of our approach to support the analysis of domain experts in archaeology and the field of CH in general.

⁷<https://www.meshlab.net/>

4.4.1 Concept

Our concept supports the similarity search over a database of images with an incomplete 3D object as an example input. Due to the different modalities of search space and reference (images as opposed to incomplete 3D models), feature-based methods are not directly applicable, but several processing steps are necessary to transfer them to a common base modality. As the inference of 3D models from images is an ill-posed problem [143], the obvious choice for the base modality is images since 3D models can be easily mapped to image space using rendering techniques. FIG. 4.4 shows our proposed processing pipeline. For the reference input, a 3D model, the first step is to orient it in a 3D viewer so that the view is consistent with the standards used for archeologic photographs [53, 54, 147]. Next, a render with near-orthographic projection is generated. If the input is incomplete, the user now has the opportunity to roughly outline the missing silhouette parts through a sketch interface (SEC. 4.4.2.1). Since we found that gradient-based FDs yield the best results, we have to fill all newly-created empty areas with plausible texture. Said is accomplished with a texture synthesis technique (SEC. 4.4.2.2), which fills the areas in question with adequate texture based on the available original surface. The result is a digitally-restored query image usable for a conventional image-based similarity search (SEC. 4.4.3). Note that the objective of this query preparation is not to create a photorealistic reconstruction of the missing object parts but rather to specify the extent and shape of the complete object and the presence of various appendages in an intuitive manner. Optionally, the query can be further refined by applying a region-based weighting to the query image (SEC. 4.4.3.2). This allows the user to influence the search by shifting the focus to essential parts of the geometry or disregarding parts that are irrelevant or misleading for the given input. Dedicated visual analytics tools (SEC. 4.4.3.2) allow a user to assess which parts of an object were decisive for the current results. The gained insight can be used to refine the query weighting iteratively.

For the target objects (search space), we consider a heterogeneous set of archaeological photographs (SEC. 2.2.4). They also require additional preprocessing (SEC. 3.2.1) to exhibit the same characteristics as the reconstructed query.

4.4.2 Query Preparation

This part of the proposed pipeline takes the original (incomplete) 3D model as input and outputs the reconstructed and normalized query image. First of all, the model has to be properly aligned by the user before a snapshot is rendered. To this end, our prototype offers a plain 3D viewer using a near-orthographic projection. Together with collaborating domain researchers, it was established that the archaeological depictions come closer to an orthographic projection than to a perspective one, as the photographs are typically taken from far away with a very narrow field-of-view angle [53, 54, 147]. *I.e.*, after loading the model of the object to search for, the first step is to orient the fragment properly. The goal is that the displayed model is oriented such as the complete object would appear in a photograph. Standardized pictures show the objects

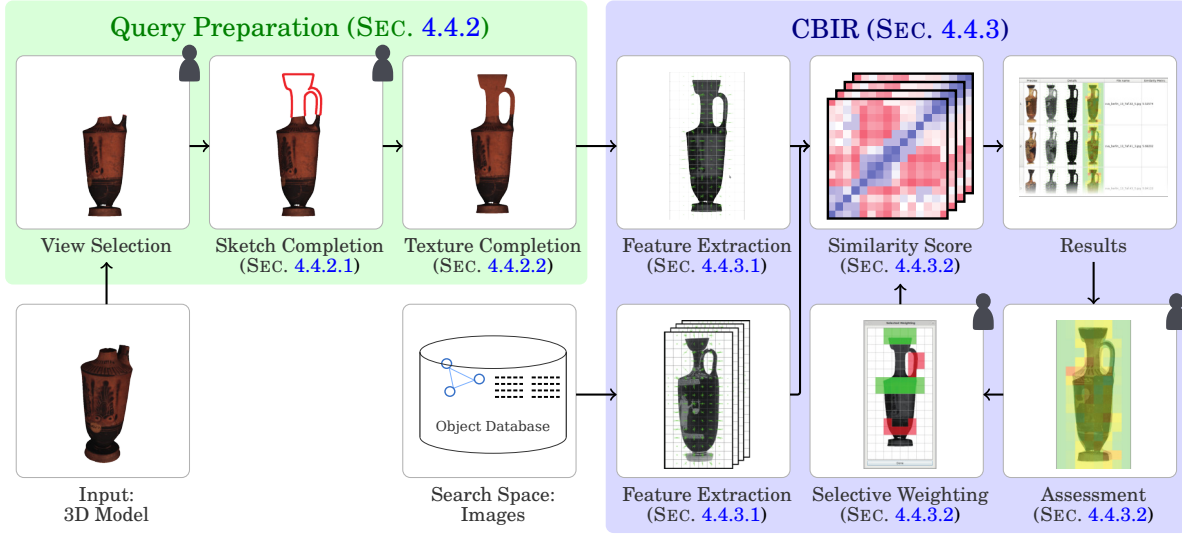


FIGURE 4.4. The concept for implementing shape-based search using hybrid 3D+Sketch queries. The sketch-completed input object is transferred to image space and compared to the target images resulting in a ranked list of results. Optionally, the results can be iteratively evaluated and refined by region-based selective weighting. Steps requiring user feedback are indicated with a characteristic icon.

standing upright from canonical angles. While approaches exist for estimating the rotation axis from a sherd of a pottery artifact [24, 195–199], an automatic pose estimation is error-prone and yields only an axial orientation. Hence, orienting the sherd *w.r.t.* the camera would still require human intervention. That is, we completely outsource the task of finding the optimal view to the user, who has to rotate the model accordingly. Once satisfied, a high-resolution snapshot is taken, and the viewer seamlessly transfers into a sketch interface. There, missing parts can be outlined by sketch (SEC. 4.4.2.1) before the thereof newly generated surface parts are filled with plausible texture (SEC. 4.4.2.2).

4.4.2.1 Contour Completion of Query Object with Sketch

Sketch lines outlining the extent of the complete vessel are drawn on a 2D overlay, superimposed onto the rendered object. These sketch lines do not need to be perfect but should only serve as a rough orientation. Yet, as we require closed surface boundaries for the subsequent inpainting step, the sketch interface implements a means to compensate for holes in the contour. We consider three error cases (FIG. 4.5): ① gaps between the sketched lines and the remaining geometry; ② gaps between adjacent sketch lines; and ③ overlappings resulting from sketch lines. These issues are handled as follows. In the case of ①, the sketched line is connected to the closest point on the geometry. In the case of ②, sketched lines are closed by connecting their closest endpoints. And case ③ is addressed by connecting the lines at their intersection point and removing the parts which are not connected to the object or other lines. The sketch lines are stored as paths,

but once the sketching process is finished, the sketch lines are drawn on the object render as thin border lines. This result is then passed on further down the pipeline.

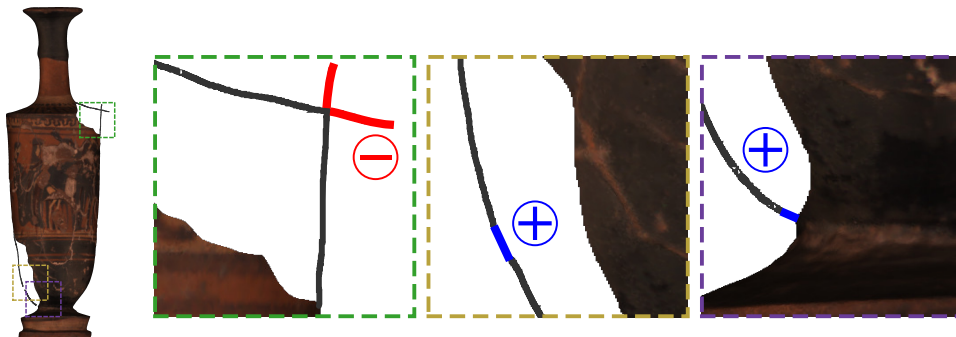


FIGURE 4.5. Completing a 2D shape by hand with a freehand sketch can result in various deviations from an optimal watertight silhouette. *I.e.*, intersections between sketch lines result in superfluous appendices (green), adjacent sketch lines having gaps in-between (yellow), and gaps between the sketch lines and the shape to be completed. While the first error case requires the removal of lines, additional lines have to be added for the latter two in order to obtain a clean and watertight silhouette.

4.4.2.2 Texture Completion

For the CBIR, empty surface areas need to be filled with texture in a manner that could plausibly resemble the original texture. The objective is to alter these regions so that they do not interfere with gradient-based FDs. Hence, flood-filling the whole region is not expedient, as it would introduce a strong gradient near the fracture line and an implausible null gradient everywhere else. Instead, the aim is to achieve similar color and coarseness properties as in the original texture.

Filling the missing parts of an image based on its remaining information is commonly referred to as *Image Inpainting* and has been an intensively researched topic for many years. The objective of such a process is to optimally estimate a subset of an image’s pixels, given by a target mask $\mathcal{T}$. To this end, usually, all image pixels (for some approaches even additional images [200, 201]) minus $\mathcal{T}$ are taken as the source. In our case, however, we use exclusively the pixels showing the displayed object as the source. A mask $\mathcal{S}$ for these can be directly obtained from the rendering process. $\mathcal{T}$ is also easily inferred by flood-filling the background combined with fundamental set operations.

The number of inpainting methods is vast and can be split into several categories, each of them with its respective advantages and disadvantages [202]. For our application, we test two different types of methods. First, the *Exemplar-based Texture Synthesis* algorithm by Wei and Levoy [203], which is combined with a FD loosely based on the Rotation-Invariant Feature Transform (RIFT) descriptor by Lazebnik and Schmid [204]. The basic idea is to compute a pixel descriptor for each pixel of $\mathcal{S}$ based on the color histogram of its neighborhood. $\mathcal{T}$ is then filled linearly from top to bottom and left to right. For each pixel to fill, we similarly compute the FV

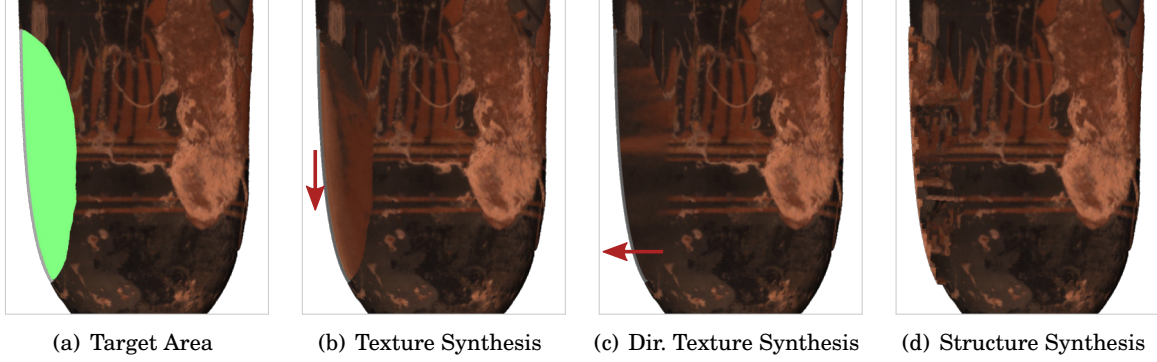


FIGURE 4.6. (a): The input plus sketch completion defines the area to be filled (green). The inpainting with the exemplar-based texture synthesis [203] (b) can be combined with a user-provided filling direction for more organic gradients following the original texture of the object (c). (d): With an exemplar-based structure synthesis [205], the area is filled iteratively patch-wise.

and use it to determine the L_2 distance to all pixels of S . From those with a distance below an empirically determined threshold, one is selected at random, and its color is applied to the current pixel in question. If no neighbor within this threshold is found, the color of the pixel with the lowest distance is taken. Finally, a median filter is applied to remove noise possibly caused by the sampling. FIG. 4.6(b) shows an example result obtained with this approach for the input in FIG. 4.6(a). It can be seen that this inpainting method introduces a strong vertical directivity, and thereby unwanted gradients, to the filled regions. Instead, the desired behavior is that the inpainting originates at the fracture line and propagates toward the sketched outline. To this end, we adapt our system by a means that allows the user to specify the filling direction (FIG. 4.6(c)).

The second approach we test is the *Exemplar-based Structure Synthesis* algorithm by Criminisi *et al.* [205]. In contrast to the texture synthesis, the idea of structure synthesis is to fill the target region iteratively from the region borders with patches of a fixed size. For each iteration, weights are calculated along the filling front. Those are a combination of a certainty term, modeling the uncertainty with growing distance from the region border, and a data term, favoring areas with a strong gradient normal to the region border. The point with the maximal weight is taken as the origin for propagating the filling front. The neighborhood of this origin is used for finding a fitting patch from the source S . The strength of this approach is that gradient information is diffused into the target region, usually leading to a more organic textual reconstruction. As can be seen from the results in FIG. 4.6(d), this type of synthesis clearly has the most realistic appearance. This is to be expected as texture synthesis is known to perform poorly for the reconstruction of large areas [202]. On the other hand, the better texture obtained with structure synthesis comes at the expense of being generally more time-consuming. Yet, with computation times in the range of a few minutes for large inputs, the better results from the method by Criminisi *et al.* [205] outweigh the more efficient computation Wei and Levoy [203] for our purpose. Hence, the prior was used for all subsequent evaluations.

4.4.3 Content-based Image Retrieval

After the query preparation steps, query and target objects are present in a common base modality (image), allowing the employment of a conventional feature-based similarity search. There exists a wide range of established image features to this end, incorporating global features such as color histogram, edge histogram, coarseness [97], Color and Edge Directivity Descriptor (CEDD) [96, p.30-40], Histogram of Oriented Gradients (HoG) [94], and local features such as Scale-Invariant Feature Transform (SIFT) [95]. Global features are computed from the whole image and can be further divided into features based on color, texture, and shape. The group of local features relies on ‘significant’ points in the image, which noticeably differ from their neighborhood. Both local and global features can be aggregated into a Bag-of-Visual-Words (BoVW) [64–66]. More recently, ML-based methods – *w.r.t.* image retrieval CNNs in particular – achieve state-of-the-art performances [206].

We found that color-based features perform poorly for our specific use case. On the one hand, the search space photographs exhibit mostly indiscriminating color distributions. On the other hand, many of the (older) archaeological photographs are published exclusively in grayscale. The evaluated local features were also not able to capture the desired notion of similarity. While they are very successful for registration purposes, *i.e.*, finding one and the same object in different images, the presence of certain keypoints turns out to be insignificant for the similarity of two vessels. Also, first experiments with a general purpose CNN [100] showed that neither an off-the-shelf network is able to grasp an expert’s notion of similarity. Training a network from scratch or the application of transfer learning techniques is hindered by the availability of necessary training data. The most promising results were obtained using global shape or gradient-based features. Scale and/or rotation invariance of the descriptor is thereby no necessity as the search space images are in compliance with established conventions [53, 54, 147].

4.4.3.1 Feature Descriptors

Two FDs are evaluated in more depth: The HoG by Dalal and Triggs [94] and the Shape Contour Descriptor (SCD) by Attalla and Siy [98]. Even though they are no recently developed concepts, they are well suited for the intended application and provide a good trade-off between efficiency and complexity.

Histogram of Oriented Gradients. The HoG describes an image by an array of local gradient histograms. To this end, the input is divided into a fixed number of equal-sized blocks which are in turn subdivided into cells. For each cell, we calculate the directivity and magnitude of gradients. It is required that the inputs are normalized such that objects occupy the same areas in the compared images. We achieve said normalization with a series of fundamental image processing steps. Note that we know the extent of the displayed object in the image as we have object masks for the images in the search space (*c.f.* SEC. 3.2.1). This allows us to crop the images

to the straight bounding box of the displayed object, which is easily derived from the minima and maxima positions of the masked pixels. As it is unfavorable for the gradient computation if the boundaries of the displayed object touch the boundaries of the image, we add additional padding, amounting to 10% of the image’s width or height (whatever is larger), around the centered object in order to create additional context. The input for the HoG descriptor is required to have a fixed width and height. To this end, we again add additional padding, either to the top and bottom or left and right to meet the required aspect ratio. Note that all newly added pixels, as well as all pixels outside the object mask, are filled with uniform white color in order to introduce a strong gradient between the background and the generally darker object. Finally, we scale the image to the required dimensions. An input image size of 128×128 for bulgy objects and 64×128 pixels for slender ones proved appropriate. As most objects in the dataset belong to the latter group, we exclusively use the latter input size for all subsequent experiments. Regarding hyperparameters, we use the settings recommended by Attalla and Siy [98] with a block size of 16×16 pixels, cell width of $s_{cell} = 8$ pixels, and $r_{bins} = 9$ radial histogram bins, resulting in a FV of 3,780 floating point numbers.

Shape Contour Descriptor. The SCD is solely based on the shape of an object, requiring that the object mask is correctly extracted. From this mask, the contour can be obtained efficiently with a topological connectivity analysis [207]. In our case, only the extreme outer contours are extracted. The concept of the SCD is to split the contour into n equal-sized chords, starting from the point with the largest distance from the shape’s center. In our experiments, $n = 100$ proved to be a good trade-off between detail and efficiency. For each chord, the descriptor stores three properties: (i) the angle between the current chord and the edge from the chord’s start point to the shape’s centroid; (ii) the distance between the chord’s start point and the centroid; and (iii) the smoothness, given by the ratio of chord length to the arc length of the contour. All three features are normalized using the max values of the input shape. Due to this normalization, the descriptor is invariant to scale, rotation and translation.

4.4.3.2 Distance Metric and Relevance Feedback

Either FD yields a FV $\mathbf{v} \in \mathbb{R}^d$ of fixed dimensionality d for any input. The similarity of two inputs a, b is assumed to be relative to the similarity of their respective FVs $\mathbf{v}_a, \mathbf{v}_b$, which is determined using a distance metric. A family of commonly used distance functions is the *Minkowski* distance

$$(4.1) \quad L_s(\mathbf{v}_a, \mathbf{v}_b) = \left(\sum_{i \in [0..d)} |\mathbf{v}_a[i] - \mathbf{v}_b[i]|^s \right)^{1/s},$$

with $s \geq 1$. Prominent representatives of this distance function are the *Manhattan* distance (L_1) and the *Euclidean* distance (L_2). For our experiments, we used the latter. Note that in our approach, the used FD and distance metrics are readily exchangeable. Integration of other features using feature combination or selection methods would be straightforward. The principle

of a feature-based similarity search system is to compute the distances between the query and all objects in the search space, giving a ranked list of similar objects with ascending distances to the query.

We found that – in contrast to translation and scale – reflection-invariance is of relevance for our application, as objects are photographed from different sides. Specifically for objects exhibiting a degree of asymmetry, *e.g.*, vases with a single handle, at least one photograph of an archaeological documentation shows the vessel such that the handle is either on the left or the right side. While in some fascicules this ambiguity is standardized, *e.g.*, by showing all objects with the handle on the right side [137, 140], this is not the rule in general. Yet, this discrepancy is entirely irrelevant to similarity considerations. Surprisingly, the invariance *w.r.t.* horizontally mirrored inputs is not considered in most state-of-the-art feature descriptors [208]. Reflection-invariant adaptations have been researched, such as the Mirror Reflection Invariant HoG (MI-HoG) by Kanezaki *et al.* [209] for the HoG. However, we found that this degree of sophistication is not required for our application. Instead, we ensure reflection-invariance on a general level by also extracting a FV $\mathbf{v}_a^*$ from the horizontally flipped input image and applying the adapted distance function $L_s^M = \min\{L_s(\mathbf{v}_a, \mathbf{v}_b), L_s(\mathbf{v}_a^*, \mathbf{v}_b)\}$.

Selective Weighting If assumptions regarding the significance of individual FV elements can be made, the *Minkowski* distance can be extended with a weights vector $\mathbf{w} \in \mathbb{R}^d$ [87] (as a multiplicative factor inside the sum in Eq. (4.1)), governing the influence of individual elements. Such assumptions can be based either based on heuristics (if present) or inferred directly from a *Relevance Feedback*. The concept of relevance feedback (also referred to as *user feedback* [170, p.9]) was introduced to content-based multimedia retrieval in the 1990s and has proven to improve the performance of retrieval systems significantly [210]. With the approach by Elad *et al.* [211], for example, a user can label the results of a feature-based 3DORas either ‘relevant’ or ‘irrelevant’, which directly influences the distance metric for the next run. In our case, we do not let the user label result items but specify the relevance of image regions directly, as we assume that the displayed objects are congruent due to our normalization effort. Following the idea of Rocchio [212], this should lead to an increasingly precise query vector with every feedback cycle.

W.r.t. ancient Greek pottery, things like the presence of a handle may be irrelevant, while a domain researcher might want to put the focus on other query attributes, *e.g.*, the shape of a vase’s shoulder or spout. To this end, we define three distinct tiers of importance, which are assigned to each pixel of the query image. Their respective impact is determined empirically with $\kappa_{max} = 5.0$ for ‘important’ pixels, $\kappa_{min} = 0$ for ‘irrelevant’ pixels, and $\kappa_{default} = 1.0$ for all other pixels. For assigning these tiers to image regions, our prototype implementation provides a selection interface, allowing us to mask important and irrelevant regions intuitively. To this end, we implement a superimposed selection grid for exclusive weighting (FIG. 4.7, left), as well as a brush tool with a Gaussian drop-off kernel, supporting a freehand selection (FIG. 4.7, right). The resulting user weighting is saved to an intensity matrix K of equal size to the input image.

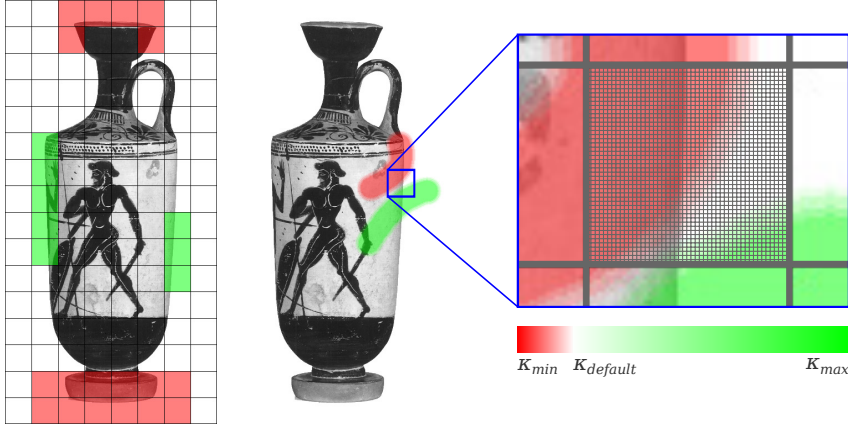


FIGURE 4.7. The importance (green) or irrelevance (red) of object attributes for the retrieval can be specified via a selection grid (left) or a freehand brush (right). For the latter smooth boundaries are achieved using a Gaussian drop-off kernel.

The mapping of a pixel weight $\kappa_{ij} \in K$, of this matrix's i -th row and j -th column, to the weights vector $\mathbf{w}$ depends on the applied feature descriptor. For the cellular HoG, we first of all, compute a cell-wise mean weight, given by

$$(4.2) \quad \bar{w}_{kl} = \frac{1}{s_{cell}^2} \sum_{\substack{i \in [k \cdot s_{cell} .. (k+1) \cdot s_{cell}] \\ j \in [l \cdot s_{cell} .. (l+1) \cdot s_{cell}]}} \kappa_{ij}$$

for the cell $\langle k, l \rangle$. We flatten these weights in a similar fashion to the elements of the FV, but instead of the gradient magnitudes, we use the mean weights. Note that r_{bins} elements are added per cell to match the length and order of the FV. With the SCD the situation is more straightforward, as a FV element already corresponds to a specific pixel in the input image. This correspondence is used to directly assign the weights κ_{ij} to the FV.

Quality Assessment. We pair the relevance feedback with a tailored visualization, revealing to the user which object parts are the most influential for a given retrieval. To this end, we illustrate – on a per-result-item basis – which regions are remarkably similar or dissimilar to the given query. This helps the user to understand why a particular object is at its specific position within the similarity ranking. The appearance of this visualization depends on the applied FD.

For the HoG, we can link the element-wise distances (Eq. (4.1)) to cells of the image, by inverting the aforementioned FV flattening. We average those distances in a cell-wise manner, giving a ‘similarity grid’, which we visualize using a semi-transparent heat map superposed to the input image. Our implementation supports a traffic light (green to red) color scheme showing the absolute differences (FIG. 4.8, left) and a seismic color scheme showing the signed differences (FIG. 4.8, middle). With this heatmap view, a user can perceive local similarities and dissimilarities between a pair of images. For the SCD, the FV values can be attributed to specific points of the object contour. Thus, the resulting n -dimensional vector of distances is directly used

to illustrate the similarities of contour point pairs, as shown in FIG. 4.8, right. The stages of relevance feedback and quality assessment (*c.f.* FIG. 4.4) can be repeated indefinitely.

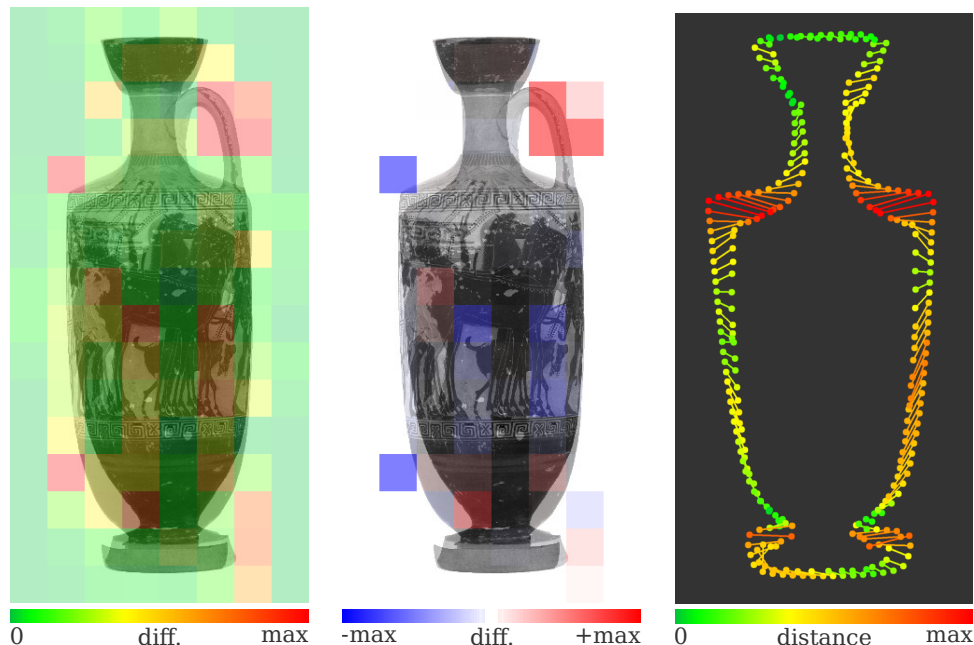


FIGURE 4.8. Local similarities according to the HoG are revealed with a traffic light color scheme for absolute differences (left) or a seismic color scheme for signed differences (middle). For the SCD, the object silhouettes are overlaid, and corresponding contour points are connected with a line colored according to their respective similarities.

4.4.4 Results

In the following, we evaluate our pipeline qualitatively (SEC. 4.4.4.1) with a visual analysis of the results for a few well-selected meaningful examples, as well as quantitatively (SEC. 4.4.4.2) through synthetically generated fractures and sketches. Prior to that, we briefly discuss the implementation details of the prototype retrieval system we employ for the experiments as well as the data used in the experiments.

Implementation. The proposed pipeline is implemented in a functional prototype (FIG. 4.9). All steps can be conducted seamlessly with this purposely built software. For the sketch-completion part (SEC. 4.4.2.1), a 3D viewer is provided, which is capable of loading textured 3D models as PLY files and supports sketching directly on top of the 3D view. The rendered result can be exported for the subsequent (optional) texture completion step (SEC. 4.4.2.2), which is conducted in another window. The program is implemented in C++ with the GUI based on Qt5⁸. We also heavily rely on the OpenCV⁹ Open Source Computer Vision library, both for tasks of

⁸<https://www.qt.io/>

⁹<https://opencv.org/>

image preprocessing and feature extraction. Note, however, that the individual stages of our retrieval pipeline in FIG. 4.4 are independent of one another. Hence, intermediate results can be fed into and extracted from the program via the file system. *I.e.*, one could use, for example, arbitrary 3D and graphics editors to generate the input for the texture inpainting step.

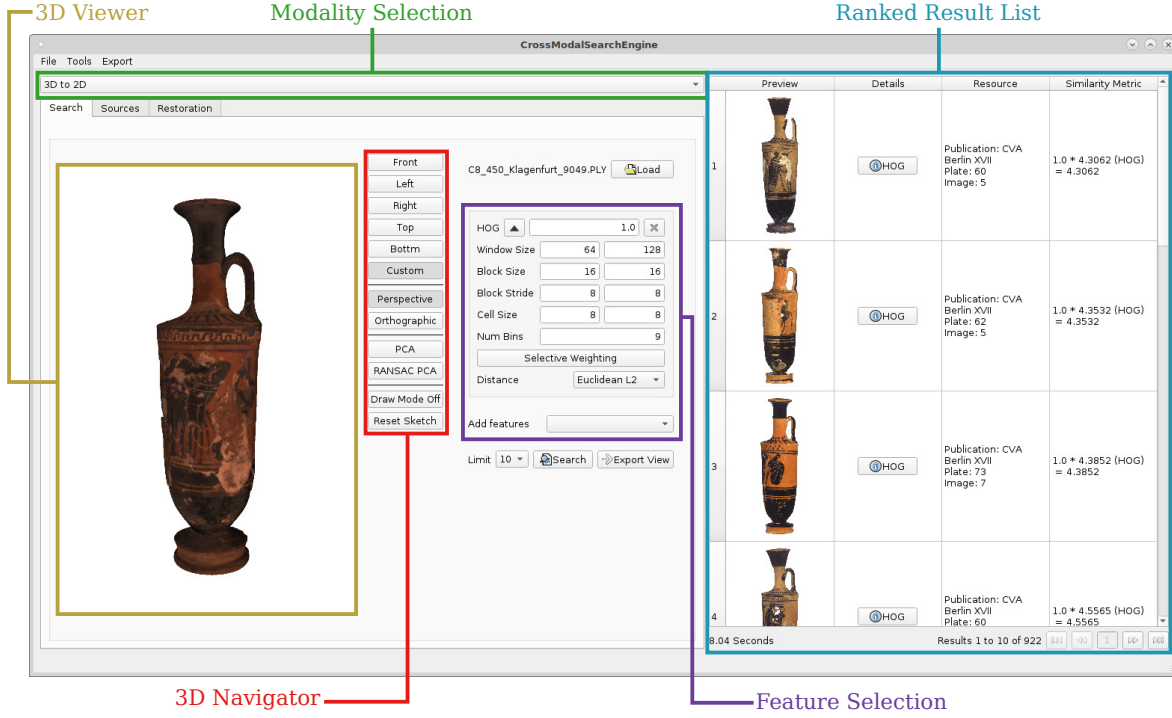


FIGURE 4.9. GUI of the prototype used for sketch-aided retrieval. Note, that the sketching can be done directly in the 3D viewer. Once a user starts to sketch, the 3D model cannot be reoriented any longer.

Data Basis. Regarding 3D models for queries, we decide on a set of seven Attic black-figured lekythoi, courtesy of the LMK (SEC. 2.2.5), from the first half of the 5th century BCE, a prevalent vessel type in Greece for that period. Additionally, we experiment with Attic Geometric pottery from the KHM, exhibiting vastly different vessel shapes and painting styles. All of the models are subjected to synthetic fracturing, which removes parts of their geometry, to mimic the effects of poor conservation status, *e.g.*, with a broken handle, mouth, *etc.* Different degrees of fragmentation are created this way. As a search space, we use 3,340 images we extracted from the fascicules CVA Karlsruhe 1; CVA München 3; CVA Kassel 1; CVA Louvre 16; CVA Würzburg 1; CVA Tübingen 2; CVA Athens 5; CVA Metropolitan Museum 5; CVA Berlin 10; CVA Oxford 4; CVA British Museum 11; CVA Berlin 13; CVA Bonn 4; and CVA Berlin 17 (*c.f.* TAB. 2.2), showing mostly lekythoi shapes and Geometric pottery.

4.4.4.1 Qualitative Evaluation

We separately assess the benefit of sketch completion and feature weighting for improving the search result in a retrieval system, with a particular focus on domain tasks in archaeological object comparison. To this end, we use one specific vessel, the black-figure lekythos *KF9049* (FIG. 4.10, upper left corner). It is agreed upon modern by domain researchers that the main characteristic for the attribution of this type of vase is the details of their shape [13]. An unfractured example was picked deliberately, as it allows us both the evaluation of the reconstruction as well as the comparison of the retrieval to a ground truth. Yet, as the pipeline expects an incomplete input, we manually remove the lekythos’s handle and spout (FIG. 4.10a), well known weak spots of the geometry. This results in a reasonable amount of missing geometry while mimicking a characteristically placed fracture line, similarly exhibited by a large number of de facto fractured objects. We compare the performance of the SCD, focusing on matching the object silhouettes to the HoG, using local gradient information.

Reference Ranking. As a ground truth for the retrieval, archaeologists determined the 15 most similar target objects according to domain consideration from all reference objects in the search space. This ranking was established using a holistic approach by considering the vessel shape and the style of painting in toto. That is, they differentiate the objects based on the main Greek painting styles (*i.e.*, black-figured, red-figured), and then perform an exclusion of specific shapes (*e.g.*, bulgy shapes) that are dissimilar to the query. The remaining target objects are ranked by looking for analogies to the reference, a common approach in archeology. This results in the 15 objects, documented by 30 images (two per object) in FIG. 4.10.

Retrieval with Fractured Input FIG. 4.10a displays the top 30 images of the ranked result set retrieved using the incomplete input object. The upper row shows the retrieval results using the SCD while the lower row shows the HoG-based results. It can be seen that both descriptors are able to retrieve images visually similar to the query image, ranking equally thin-shaped vases first. The rotation-invariance of the SCD also yields some outlier results of flat vessels having a similar, but differently oriented, silhouette. *W.r.t.* to the reference ranking, we can observe that none of these best-ranked references is present in the reference set, *i.e.*, this result does not reflect the experts’ notion of similarity. However, this is an expected result, as the applied FDs cannot compensate for missing geometry. Hence, the top results are mostly comprised of similarly broken vases.

Sketch-Aided Retrieval FIG. 4.10b illustrates the ranked results based on sketch-completed query objects. Here, we use the rendered object with the sketched silhouette as a thin black line but without any additional texture completion. A visual inspection of the top 30 results shows a much stronger resemblance to the reference results for both descriptors. Numerically



FIGURE 4.10. Ranked result sets obtained from different retrievals of *KF9049*, using an incomplete query object (a), the sketch-completed input (b), additional texture inpainting (c), and additional selective weighting (d). The top row shows the 30 most similar entries from the given search space, ranked by similarity according to archaeological knowledge. If a result object is present in this reference set (match), the fill level of the blue background indicates the accuracy (EQ. (4.4)) of its position compared to this reference ranking. The numbers on the right indicate both the number of matches between retrieval and reference sets as well as the distance between them according to EQ. (4.3).

this improvement manifests itself with 15 and 11 of the top 30 matches overlapping with the reference for the HoG and SCD, respectively.

Besides a mere matching count, we also investigate the similarity of the rankings in more depth. To this end, we measure each matching image's ranking deviation from its ranking in the reference set. Specifically, we determine the quality of a retrieval with the result vector $\mathbf{v}_{ret}$ w.r.t. a n_{ref} -dimensional reference result vector $\mathbf{v}_{ref} \in \mathbb{R}^{n_{ref}}$ by the normalized mean *Manhattan* distance

$$(4.3) \quad D(\mathbf{v}_{ret}, \mathbf{v}_{ref}) = \frac{1}{n_{ref}^2} \sum_{i \in [0..n_{ref})} D_i(\mathbf{v}_{ret}[i], \mathbf{v}_{ref}),$$

with

$$(4.4) \quad D_i(v, \mathbf{v}_{ref}) = \begin{cases} |i - j| & , \text{ if } \exists j \in [0..n_{ref}) : v = \mathbf{v}_{ref}[j] \\ n_{ref} - 1 & , \text{ else,} \end{cases}$$

as the accuracy of a specific result item i . In FIG. 4.10, these accuracies are indicated by the filling level of the blue bar in the background of matching vessel images, with a completely filled bar suggesting a perfect correlation between retrieval and reference position ($D_i = 0$), while a completely empty bar, contrarily, suggesting the largest possible offset ($D_i = n_{ref} - 1$). The overall distances D are given for each query, together with the hit count, on the right hand side of FIG. 4.10. If we compare the HoG to the SCD results from an archaeological point of view, we observe that the HoG performs better. We attribute this to the fact that the reference set is – like the query object – comprised of black-figured lekythoi. The color agnostic SCD ranks several red-figure lekythoi with similar shape proportion also as top hits.

Texture Inpainting. With additional texture inpainting of the sketch-completed query (FIG. 4.10c), the HoG-based result improves significantly, giving 22 instead of 15 matches and a decreased distance of $D = 0.51$. Note that this inpainting step does not affect the retrieval using the SCD, as this descriptor incorporates the silhouette information solely.

Selective Weighting. The quality of the results can be further refined by incorporating expert knowledge, thereby adjusting the weighting for different regions. In the case of the lekythos *KF9049*, archaeologists experimenting with our prototype implementation put the focus on regions containing original texture and shape and disregard the importance of the presence of a handle. This reasoning is based on their domain experience that similarity relations of lekythoi have to be detected in vessel profiles, whereas attachments such as handles are less significant [213]. FIG. 4.10d shows the corresponding weighting (left) and the ranking results obtained incorporating these weights. We observe a further improvement of the number of matches of 13 and 23 out of 30, as well as a yet smaller overall distance of $D = 0.70$ and 0.41 for SCD and HoG, respectively. For the following experiments, the HoG was used exclusively as it generally yields better results. We attribute this to the fact that shape information alone is not sufficient to model an expert’s notion of similarity. By relying on gradient information, the HoG is not only capturing an object’s shape but also indirectly colorimetric information.

Reflection Invariance. While the images in the reference set in FIG. 4.10 mostly show lekythoi with a single handle which is on the right side, the side from which an object is photographed may vary from fascicule to fascicule or even within a single publication. Still, the object’s orientation should not influence the similarity computation. Hence the retrieval must be invariant *w.r.t.* a target appearing in horizontally mirrored fashion *w.r.t.* to the query. For evaluating the reflection invariance of our search, we choose the Attic Geometric pitcher *KHM IV*

1 (FIG. 4.11, left). The object belongs to a specific group of pitchers exhibiting globular bodies with vertical gadroons and a broad concave neck, all dated to Late Geometric Ib (c. 750-740 BCE). Based on similarities in vessel shape and vase painting, three further pitchers are attributed to this workshop or, more specifically, to the same painter/potter [214]. Those are *Karlsruhe B 2680* (a), *Oxford 1894.13* (b), and *British Museum GR 1878,0812.8* (c).

Same as with the lekythos, we introduce a reasonable amount of fragmentation by artificially removing the handle (FIG. 4.11, top row, left), resembling a realistic state of incompleteness. Even with the retrieval conducted with the incomplete object, our system is able to detect one of the similar pitchers: *Oxford 1894.13* (b), highlighted with dark blue background in the top row of FIG. 4.11. The matches on rank 2 and 4 (highlighted in light blue) are other pitchers with similar proportions of the vessel’s shape, *Louvre CA 1940* (e) and *Tübingen 6214* (g). When applying our sketch-aided retrieval, but without our adaption for reflection invariance (SEC. 4.4.3.2), *Karlsruhe B 2680* (a), the best preserved similar pitcher, is ranked in first place (FIG. 4.11, 2nd row). Additional pitchers are on ranks 2, 3, 6, 9, and 10 with *Tübingen 6214* (g), *Berlin 31044* (h), *Würzburg H 5171* (d), *Louvre CA 1940* (e) and *Tübingen S./10 1088* (f). According to the primary publication of *KHM IV 1* [214], the Würzburg pitcher is very close to the Vienna pitcher regarding the vessels’ shape and proportions. Oxford, which was previously ranked in first place, is no longer within the top 10 since it is photographed only from the front, the back, and with the handle on the right side. With a reflection invariant search (FIG. 4.11, 3rd row), all of the three pitchers attributed to the same workshop, *Karlsruhe B 2680* (a), *Oxford 1894.13* (b), and *British Museum GR 1878,0812.8* (c) are ranked within the top 10 together with 4 other pitchers already mentioned above. These are all pitchers from the Late Geometric Ib/IIa period (3rd quarter of the 8th century BCE), relevant results albeit being from different workshops as the *KHM IV 1*.

4.4.4.2 Quantitative Evaluation

Due to several steps of our proposed pipeline requiring user interaction, a manual evaluation with a large number of queries is not feasible. Therefore, we evaluate the query completion (SEC. 4.4.2) and the CBIR (SEC. 4.4.3) components of our system unsupervised by generating a large number of synthetic queries exhibiting different degrees of fragmentation. We simulate the user input for sketch completion with varying levels of sketching proficiency. Two different scenarios are evaluated: the retrieval of *KF9049*, the lekythos shown in FIG. 4.10, and the retrieval of arbitrary vessels picked at random from our search space.

Synthetic Fracturing We aim to introduce a level of fragmentation to the depicted objects, resembling the effects of wear and decay experienced by de-facto broken artifacts. The basis for our experiments at this point is not 3D scans but the extracted photographs. Consequently, physics-driven fracturing methods, which typically require 3D models [215], are not applicable. A common method to generate plausible fracture patterns for fragile objects is to use a Voronoi tessellation of

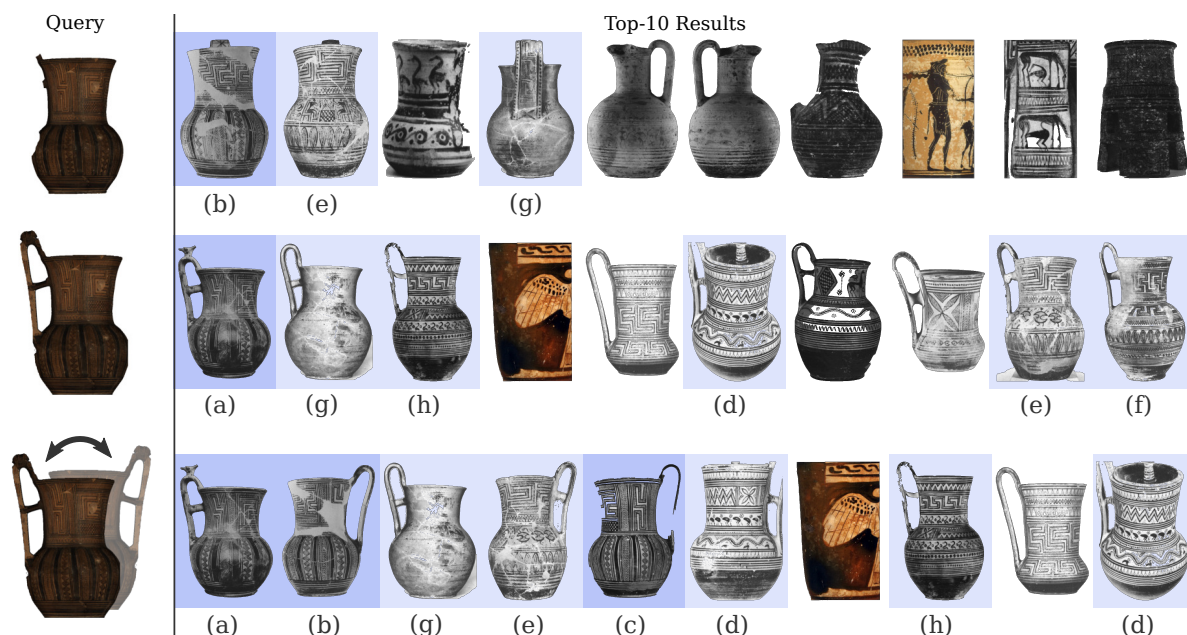


FIGURE 4.11. Retrieval results for the incomplete geometric pitcher *KHM IV 1* with the untreated incomplete query (top row), the sketch-completed and inpainted query (middle row), and the reflection-invariant distance metric (bottom row).

the object's surface. We resort to a similar approach as proposed by Banterle *et al.* [91] by cutting through a two-dimensional Voronoi tessellation superimposed on the image. The Voronoi seed points are placed normally distributed around a cutting curve function f , splitting the tessellation into two disjoint subsets. To generate a cut exhibiting a high amount of randomness and jitter, f is defined by a series of 8 superimposed sinusoidal functions $f(x') = \sum_{i \in [0..8)} m_i \sin(a_i x')$, with different random amplitudes m_i and frequencies a_i . x' and y' are the canonical basis of a coordinate system rotated by an arbitrary angle $\gamma \in [-\pi, \pi)$ w.r.t. the image's coordinate system. The system's origin P is placed at random within the subset Ω of image pixels belonging to the displayed object (FIG. 4.12, left).

A total of 100 Voronoi points $\langle x'_s, y'_s \rangle$ are seeded around $f(x')$ with coordinates $x'_s \sim \mathcal{N}_x(0, 0.15h)$, $y'_s \sim \mathcal{N}_y(f(x'), 0.1h)$, with h denoting the image height, resulting in a high cell density in the proximity of the cut. W.r.t. the arbitrary tessellation in FIG. 4.12, right, $f(x')$ divides the resulting Voronoi cells into three sets: (i) fracture cells M_f with all pixels above the curve (white); (ii) object cells M_o with all pixels below f (blue); and (iii) cells M_i intersecting the curve (orange). The final image fracturing is performed by setting all pixels $\Psi = \Omega \cap M_f$ to white color. The fraction of removed object pixels $|\Psi|/|\Omega|$ (fragmentation rate) is limited to an interval of $[0.05, 0.95]$ by discarding cuts, resulting in less or more fragmentation.

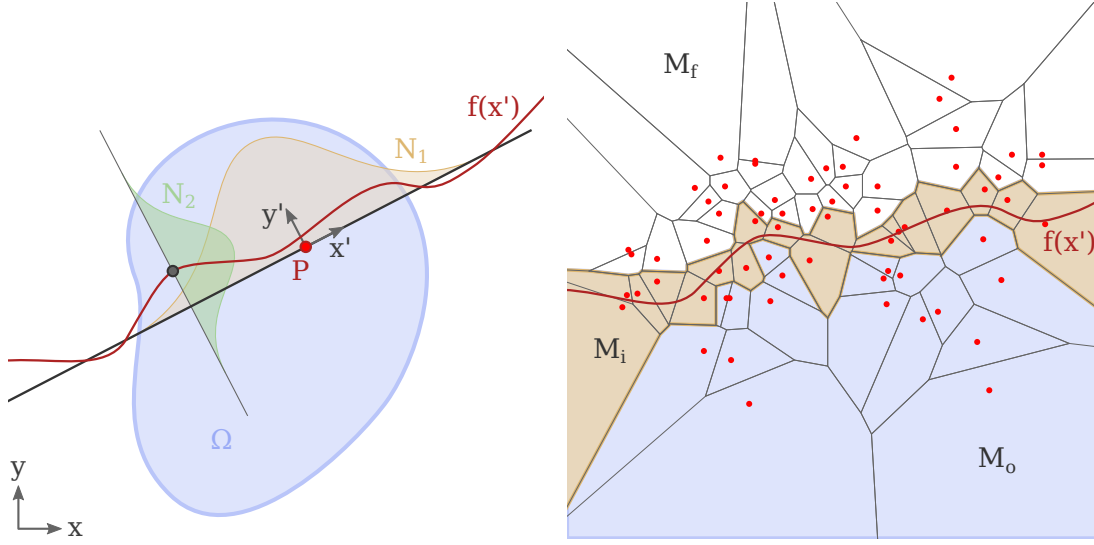


FIGURE 4.12. The fracturing is based on a cutting function $f(x')$, originating at a random point P within the object pixels Ω (left). Points placed with distribution $\mathcal{N}_1$ and $\mathcal{N}_2$ serve as seeds for a Voronoi tessellation (right).

Sketch-line Distortion. For the unsupervised evaluation system, we also have to mock the contour completion, otherwise contributed by the user. To this end, we synthetically generate sketch lines exhibiting the characteristics of user sketches with different levels of accuracy and distortion. The basis for such a sketch line is the outer contour $\delta\Psi$ of the removed object part Ψ . As $\delta\Psi$ represents a ‘perfect’ sketch of the missing silhouette, it is subjected to a series of transformations, aiming to mimic the traits of a human-created sketch. We identified three characteristics of a user-created sketch which we approximate artificially:

1. Sketch-lines are generally smooth continuous strokes without many scribbles. In contrast to that, the contours of the extracted object mask exhibit high-frequency components, either due to artifacts of the segmentation process or due to the abrasive surfaces of the artifact itself. As a countermeasure, a curve smoothing based on the concept by Mokhtarian and Abbassi [216] is applied, as indicated on the left side top of FIG. 4.13.
2. Strong curvatures are often drawn using two intersecting lines (FIG. 4.13, left bottom). We approximate said behavior by first detecting high curvature points along the contour using the algorithm by Rosenfeld and Johnston [217]. After discarding points with a curvature above a threshold angle of $\pi/3$ and a subsequent Non-Maximum Suppression (ALGORITHM 1), the remaining points are used to split $\delta\Psi$ into segments which are then smoothed independently. As a result, we get a contour exhibiting both smooth curvatures as well as sharp bends, resembling points where two sketch lines meet.
3. We experienced that even for experts, it can be hard to accurately sketch the exact proportions of the once complete object. Let $\Phi = \Omega \setminus \Psi$ be the set of pixels showing the remainder

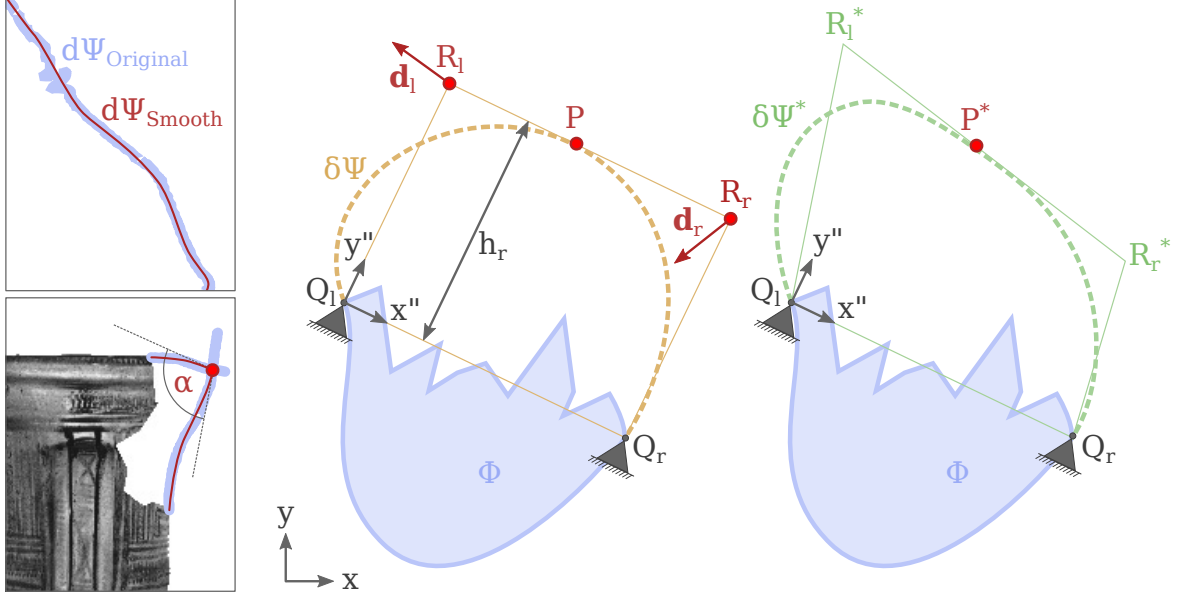


FIGURE 4.13. Left: The sketch lines are subjected to a smoothing (top) while strong curvatures are preserved using an angle criterion (bottom). Middle/Right: The linear distortion of sketch lines is governed by a control rectangle Q_l, Q_r, R_l , and R_r .

of the fragmented object, and Q_l, Q_r its intersection points with the contour $\delta\Psi$ (FIG. 4.13, middle). We observe that sketch distortions aggravate with increasing distance from Φ , and are almost zero near the bases Q_l, Q_r . We simulate this error by applying a random linear distortion that similarly increases with the distance from Φ . To this end, we define a control rectangle given by the base points Q_l, Q_r , and two reference points R_l, R_r at a distance h_R . h_R is defined as the height of the straight bounding box of $\delta\Psi$ in the coordinate system with canonical basis x'', y'' . Applying random displacement vectors $\mathbf{d}_l, \mathbf{d}_r$ to R_l, R_r allows us to map the linear distortion of this control rectangle to each point $P \in \delta\Psi$ via bilinear interpolation of the displacement vectors, giving the distorted point $P^* \in \delta\Psi^*$ (FIG. 4.13, right). Note that the displacements at Q_l, Q_r are $\mathbf{0}$.

As a measure of the overall distortion caused by these transformations (sketch line distortion), we compute the normalized mean Euclidean distance between the original contour $\delta\Psi$ and the distorted contour $\delta\Psi^*$.

Experiment Setup. We quantitatively evaluate the performance of the CBIR (SEC. 4.4.3) based on (i) the fractured image; (ii) the fractured image with distorted sketch lines; and (iii) the thereof reconstructed (inpainted) image.

KF9049 with Different Fractures. In the case of *KF9049* we can rely on the reference ranking introduced in the previous section (SEC. 4.4.4.1, FIG. 4.10) with $n_{ref} = 30$ elements.

We generate 200 different synthetic fractures plus corresponding distorted sketch lines. The normalized mean distance of the result vectors to the reference (EQ. (4.3)) is computed for all examples and for each of the fractured, sketch-completed, and inpainted queries (FIG. 4.14). Each marker encodes the orientation of the local x' , y' coordinate system (henceforth referred to by ‘cut plane orientation’) *w.r.t.* to the global x , y , as well as the fragmentation rate. The queries with the fractured input perform clearly worst, with the majority of the data points being close to maximum distance $D = 1.0$. Lower distances correspond with a lower fragmentation rate, indicating that the retrieval with the incomplete object worsens with increasing fragmentation rate. The figure clearly shows that adding the sketch outlines improves the results drastically, while the texture inpainting leads to an additional minor improvement. To summarize, as expected, the results generally degenerate with increasing sketch-line distortion as well as with increasing fragmentation rate. Moreover, there is no noticeable correlation between cut plane orientations and retrieval accuracies.

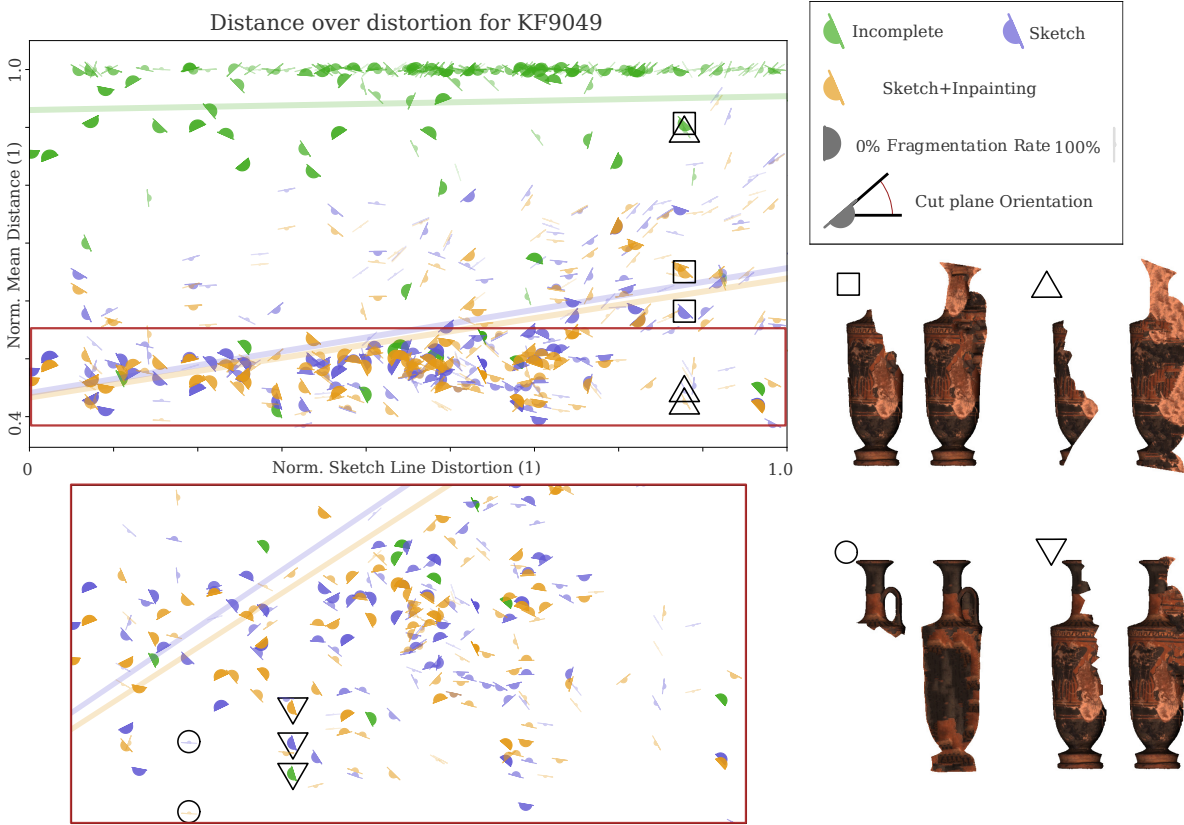


FIGURE 4.14. The normalized mean distance D over the normalized sketch-line distortion for 200 fractures plus distortions of *KF9049* with incomplete (green), sketch-completed (violet), and additionally inpainted (yellow) queries. The markers indicate the orientation of the cutting function f and the degree of fragmentation. The closeup on the lower range of the Y-axis, as well as the semi-transparent linear trend lines, reveal an additional slight performance increase with the sketch-completed+inpainted queries.

Arbitrary Vessels. For the evaluation with arbitrary queries, we manually select 55 images from our search space depicting a whole vessel. For each of them, four fractures, on average, are generated automatically, resulting in a total of 200 sample queries. For these cases, there is no reference ranking by experts available. Instead, the ranking that would result from the query with the non-fractured object is taken as a reference. Thus, we are not restricted in the number of elements of the results vectors we compare, but we could potentially compute the distance between the entire vectors. We decided to compare the first 100 elements, as it appears that there are at least 100 relevant reference images in our search space for each of the selected inputs. The mean distances over the fragmentation rate are given in FIG. 4.15.

For the incomplete queries, a similar behavior as with *KF9049* can be observed. The sketch-completed and inpainted versions perform clearly superior, but in contrast to *KF9049* the distribution appears to be wider spread. It is interesting to observe that, when considering a diverse set of objects from the search space, both sketch-completed and additionally-inpainted queries exhibit roughly similar trends as visible for *KF9049* (FIG. 4.14). We take this as an indication of the robustness of our method.

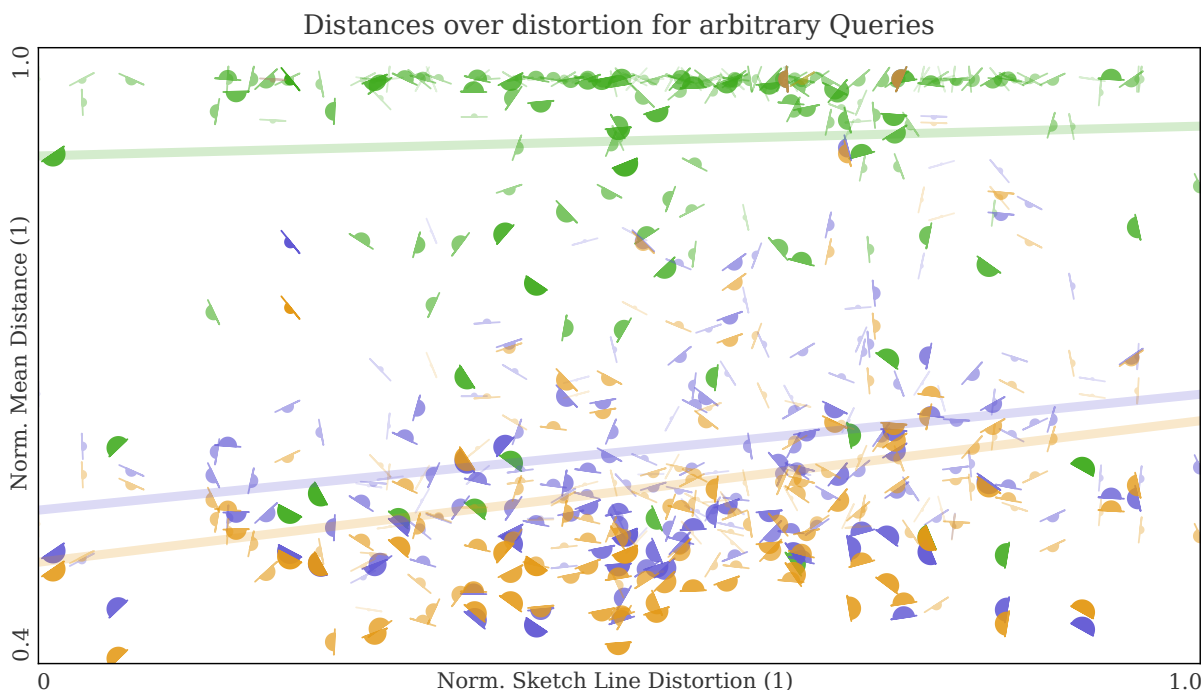


FIGURE 4.15. The normalized mean distance D over the normalized sketch line distortion for 200 experiments, based on 55 arbitrary vessel depictions and different fractures, with incomplete, sketch-completed, and inpainted queries (legend as defined in FIG. 4.14).

4.4.5 Discussion

In the following, we present several thoughts regarding the applicability of the 3D+Sketch retrieval approach (SEC. 4.4.5.1) before we also discuss some of the limitations and shortcomings (SEC. 4.4.5.2). Finally, we present a few still open research questions (SEC. 4.4.5.3).

4.4.5.1 Applicability

From our experiments, we conclude that the incorporation of user sketch-aids and relevance feedback can help to improve and refine domain-specific search tasks on incomplete shapes. With the sketch interface and the quality assessment, the domain expert can quickly fill in missing parts and visually compare the query to target objects. Sketching is an effective tool to test hypotheses, *e.g.*, whether certain object shapes or variants are present in a collection. Our experiments show that sketch-aided retrieval is generally more effective in the case of a partial search. On the one hand, a sketch allows a user to specify the overall extent of the object more intuitively than by just providing a bounding box. Said is necessary for proper rescaling, especially if parts at the top or the bottom of the object are missing. In such cases, it is the main driving factor for improvement, as shown with *KF9049* in FIG. 4.10. On the other hand, the sketch is a necessary prerequisite for texture inpainting, where it serves as a boundary. Extending the currently used HoG features with texture features could further improve the impact of the sketch-aid and inpainting on recall. With our tests, we also establish that selective weighting provides a reasonable means for domain experts to incorporate their domain knowledge. Our evaluations also demonstrate that SCD can be quite useful in cases where we want to search for similar vessel shapes independently of the painting style, *e.g.*, for a diachronic analysis of a vessel type’s shape development in different regions.

Concerning runtime, the offline preparation of the search space (image segmentation, descriptor extraction) takes several hours for all images on commodity hardware with non-optimized implementations. However, all pipeline parts pertaining to the retrieval (sketch pruning, texture completion, similarity computation) can be computed interactively, *i.e.*, within a few seconds for the whole search space with 3,340 images.

4.4.5.2 Limitations

A problem may arise if the amount of missing geometry is too large for the expert to recognize a plausible completion variant. Our experiments showed that sketch-aided search is of particular advantage in the absence of a part changing the overall shape, *e.g.*, the absence of a handle or a spout. In such cases, a definite improvement in retrieval results could be observed. As expected, sketch-completing only small or irrelevant missing parts showed no obvious advantage over standard retrieval for incomplete objects. We also note that we evaluated our sketch-aided retrieval approach on learned features using a pre-trained general-purpose neural network [100].

However, using the query object (a) and (c) shown in FIG. 4.10, those features produced only between 0 and 2 out of 30 matches. We conclude that general-purpose pre-trained networks may not produce sufficiently specialized features for the domain-specific class of objects used in our evaluations.

Two limitations are also related to the images the search space is comprised of. First, some archaeological photographs do not follow the given standards, *e.g.*, some depict a brightly illuminated object in front of a dark background (*c.f.* SEC. 3.2.1). In such cases, the automatic object segmentation fails. Thus, the SCD is not applicable and the HoG is usually severely affected by erroneous gradients in the background. Second, while our pipeline is able to cope with an incomplete query object, the same does not hold for objects in our search space with missing parts, as is the case for several of the lekythoi in FIG. 4.10a. Since the overall shape differs significantly, such objects will not be found when querying with a complete input.

4.4.5.3 Open Research Questions

Our approach currently supports low-level editing of the query shapes. Improved editing could make use of semantic-sketching techniques [218]. However, semantic sketching requires appropriate generative modeling procedures to cover the application domain, which may be expensive to obtain for many different shape types. Another idea is to guide the user while sketching in an online fashion, based on available target data, following the *ShadowDraw* idea proposed by Lee *et al.* [79]. Generally, it appears the information of how the original vessel looked like has to be present in an incomplete shape since an expert is able to complete it with a high amount of certainty. Hence, perhaps an automatically generated completion hypothesis can be presented using ML techniques if sufficient training data is available.

Regarding the CBIR, we rely on engineered (SEC. 4.4.3), as preliminary results on learned features from a general-purpose network were not able to produce meaningful results for the domain-specific data used in our experiments. As stated, the applied FD is an interchangeable part of our pipeline. Hence, other FDs can be readily added to the prototype implementation in an effort to further improve the retrieval results.

Within the presented pipeline, we investigate the retrieval over a search space of archaeological photographs. Nonetheless, the approach can be readily adapted to support the retrieval over a search space also comprising 3D models if the necessity for that comes up. To this end, the items in the search space can either be rendered (view-based), allowing to use the rest of the pipeline as it is, or, contrary to this, the incomplete query could be reconstructed in 3D [218], allowing the application of 3D-based FDs.

4.5 Conclusion

Within this chapter, we present two orthogonal retrieval approaches: NPC/W-NPC (SEC. 4.3) and 3D+Sketch (SEC. 4.4). Both combine different combinations of modalities and fill a niche in the taxonomy of approaches (TAB. 4.1). The NPC/W-NPC takes a 3D model of a complete or near complete vessel as input and supports the search over a search space of other complete or near complete 3D models using a sparse representation of the objects' profile curves. This method was developed specifically for a retrieval challenge [42] where it competed against deep-learning-based methods. Even though most of them outperform this approach, we believe the strength of this method is its generalizability, as it is applicable for arbitrary rotational solids in contrast to the other submitted techniques, which were specifically trained for the given dataset of Peruvian pottery. Another significant advantage of the NPC/W-NPC is its transparency (white box). *I.e.*, it can be unambiguously determined and visualized which parts of two objects is responsible for their similarity or dissimilarity, as well as its high efficiency due to the very compact FVs (just 50 floating-point values with our parametrization).

The newly introduced query modality 3D+Sketch requires an (incomplete) 3D model as input, which is used to query a search space comprised of images. To this end, we present an appropriate workflow and implementation, which we apply to several representative use case scenarios from the archaeological research domain. The evaluations show that the usage of user sketches and subsequent texture inpainting in order to complete and modify the 3D query object can significantly improve the search. The query modality provides the flexibility to search based on incomplete shapes in particular and to explore hypotheses of possible shapes in a target repository. While the sketch completion itself is intuitive to be applied by any user without special skills, domain knowledge of existing object shapes will lead to a better sketch and, in turn to a better retrieval result. Finally, result visualization and weight adaption support the analytical retrieval process.

To summarize, we believe that both of the presented retrieval methods represent valuable additions in their own right to the field of context-based retrieval in CH. In view of the different modalities of the available data and the collection of presented techniques, one of the core topics for future research should focus on methods that are capable bridging modality gaps, *i.e.*, *cross-modal* retrieval methods.

Part II

Texture-based Retrieval

MOTIF RETRIEVAL

Contents

5.1	Introduction	83
5.2	Related Work	85
5.3	Shape and Color-based Motif Search	86
5.3.1	Concept	87
5.3.2	Experimental Evaluation	89
5.3.3	Limitations and Discussion	93
5.4	Conclusion	94

Despite not being a research objective from the start, the similarity retrieval of pottery based on the semantic scenery depicted by their vase paintings has become a focus once it was learned that this is a relevant objective for establishing object similarity from an archaeological point of view. This chapter introduces a high-level classification of the different elements of a vase painting and presents a retrieval approach for one specific type – *i.e.*, motifs.

5.1 Introduction

As opposed to the topic of shape-based retrieval of ancient pottery has been the subject of intense research, the retrieval based solely on textual information received comparably little attention. Although some retrieval systems focus on shape, they also directly or indirectly make use of textual information, *e.g.*, in order to distinguish between black-figured and red-figured vases. Nonetheless, such systems are futile if a user wants to search for a specific depicted figure or scenery, crucial cues for the attribution to a painter or workshop [10]. To this end, specialized retrieval methods are necessary.

This chapter is based on:

- Lengauer, S., Komar, A., Labrada, A., Karl, S., Trinkl, E., Preiner, R., Bustos, B., Schreck, T., “Motif-driven retrieval of greek painted pottery”, in *Eurographics Workshop on Graphics and Cultural Heritage*, Eurographics-European Association for Computer Graphics, The Eurographics Association, 2019, pp. 91–98. DOI: [10.2312/gch.20191354](https://doi.org/10.2312/gch.20191354)

We differentiate between different manifestations of surface decorations on ancient pottery. In most cases, we observe purely colorimetric patterns resulting from paint being applied to the surface. However, surface patterns can also comprise geometric deformations of the surfaces in the form of reliefs or incisions (FIG. 5.1). Even combinations of colorimetric and geometric surface decorations are possible. We can also distinguish between ‘high-level’ paintings, *i.e.*, humanoid figures, mythological creatures, and the like (henceforth referred to by ‘motif’) and, as opposed to this, ‘ornaments’ which are characterized by their rather rudimentary shapes, which usually occur in a repetitive manner as background filling. In this chapter, we focus exclusively on motifs, while ornaments are the subject of the next chapter (CHAPTER 6).



FIGURE 5.1. Left: A Corinthian aryballos exhibiting geometric decorations in the form of incisions as well as lavish colorimetric paintings. The latter can be further divided into motifs (lions and panthers) and ornaments (rosettes and petals). Right: A botella shape from the Lurin culture features several colorimetric patterns which appear in combination with a relief revealing a face structure. (Image source for aryballos: [74]).

We present a search engine for motifs on vase paintings for the purpose of comparison and retrieval. We follow a QBE concept, meaning users are able to provide a reference image showing the motif they want to search for in a database comprised of archaeological photographs. While our considered data set is not small (hundreds of vessels), obtaining large amounts of faultless

labeled training data for supervised ML approaches is generally difficult. Therefore, we base our approach on an appropriately designed, unsupervised workflow for image segmentation, feature extraction, retrieval, and result visualization.

After a brief discussion of the related work regarding motif-driven retrieval (SEC. 5.2), we present our proposed workflow on the matter (SEC. 5.3) before we conclude the chapter with a few thoughts on the relevance and future of this kind of content-based retrieval (SEC. 5.4).

5.2 Related Work

Motif-driven retrieval falls into the very broad field of feature-based multimedia retrieval. While there have been several approaches presented pertaining to the retrieval of similar ornaments (discussed in SEC. 6.2), related work regarding the retrieval of semantically similar motifs on CH objects is very sparse, as concluded in the survey by Karl *et al.* [55].

Itskovich and Tal [219] present a partial matching scheme for reliefs with an QBE approach by detecting salient points for which they compute local FDs. Combined with a segmentation of the pattern shapes, they are able to measure the similarity between segments and reveal similar regions. They demonstrate the applicability of their approach on a series of CH objects. Two years later, Leifman and Tal [220] combine their local description approach with ML to validate the search of a given motif.

Crowley and Zisserman [221] describe an approach for the automatic annotation of gods and animals depicted on ancient Greek vessels. To this end, they leverage associated textual descriptions of vase paintings provided by the BAPD. From those, they extract predefined keywords using text-mining techniques. Consequently, they are able to link the featured gods and animals, mentioned in these descriptions, to the actual image area in which they appear. Using this knowledge, they label new images with appropriate descriptions of which figures can be seen on the painted pottery surface.

Furthermore, there are two ongoing research projects dealing with the topic of semantic analysis of ancient Greek pottery. The first project, “EGRAPHSEN. Möglichkeiten und Perspektiven der digitalen Malerzuweisung bei attischen Vasen”¹, led by Prof. Dr. Martin Langer, deals with the classification of painters and workshops of attic vases. To this end, they annotate a very large set of surface paintings, *e.g.*, figures, tools, and ornaments, but also the atomic building blocks such as eyes, hands, *etc.* These extensive annotations are used as training data for a deep neural network which should be able to decide if two vases were decorated by the same artist. In a recently published first research paper [222], they validate their proposed concept with an initial dataset comprised of 1,810 images of vase paintings attributed to 31 painters, with a total of 172,571 individual annotations. Using the learning-based classification, they obtain very mixed results with a high accuracy for object classes such as ‘mouth’ and ‘nose’, but a comparably low

¹<https://www.uni-goettingen.de/de/598165.html>

accuracy for other classes such as ‘leg’ or ‘folds’, showing that the problem is very challenging even in the presence of numerous training examples.

The second project, “Iconographics. Computational Understanding of Iconography and Narration in Visual Cultural Heritage”², lead by Prof. Dr. Corinna Reinhardt, researches the iconographic understanding and visual narration of Attic vase paintings. To this end, they use different computer vision methods, such as object detection, pose and gaze estimation, action recognition, and image captioning. They manually annotate a large set of images (9,579 images with 28,138 annotations) which they use as input for a CNN. The first results show that an mAP of roundabout 42% is possible for detecting individual items (‘sword’, ‘harpe’, ‘fish’, *etc.*) with this workflow [223].

As can be seen, most approaches rely on deep learning methods for detection. Yet, to this end, a large set of annotated training data is necessary [222] – something that is not available to us. Conversely, we aim to achieve a similar performance with the usage of engineered techniques, which has the advantages that (i) the resulting system is not tailored to a very narrow set of use cases, *e.g.*, black-figured Attic pottery; and (ii) results are easily explainable as opposed to the black-box nature of deep learning techniques. Also, most research (apart from Itskovich and Tal [219], who focus, however, exclusively on reliefs) is concerned with the classification of vase paintings and the attribution to painters/workshops. To our knowledge, there has been no system presented to date which supports querying a database for a specific motif in a QBE manner.

5.3 Shape and Color-based Motif Search

We follow a QBE design for our system, as this was established to be the best-suited variant for everyday use by archaeologists since they usually want to search for similar scenes or motifs to a given exhibit. Our approach consists of two major parts. First, generating the search space by discovering and extracting image segments, possibly corresponding to an individual motif, from domain-specific images. Second, the similarity retrieval, which we base on the contours of the extracted segments. The first part entails the preprocessing and segmentation of a potentially large set of images and therefore needs to run completely unsupervised. This poses a particular challenge for the extraction of the motifs. At this stage of the preprocessing, classifying image segments as regions of interest and non-interest is not trivial and also not desirable. Degenerated vases, poor image quality, *etc.*, often lead to bad segmentation results, which can leave classification attempts cumbersome. Therefore, instead of performing an unsupervised pre-classification based on possibly ill-defined classifiers, we include all resulting segments in our search space as possible motifs, thus preserving all the extracted information for the subsequent retrieval. Our data basis consists of a series of images depicting pottery with lavish vase paintings. The motif extraction process needs to be conducted only once for all source images. The aim is to

²<https://www.izdigital.fau.de/forschung/efi-iconographics/>

create a database of segments encompassing individual motifs. Together with its position inside the source image, the segment outlines are also used for a user-friendly result visualization.

5.3.1 Concept

From images obtained from relevant domain publications and repositories, we automatically extract a set of motifs with a tailored segmentation approach, described in SEC. 3.2.2. From those, we compute meaningful features (SEC. 5.3.1.1), which are inserted into a database, serving as the search space. A user-centered search is initiated by loading an example image and specifying the area (motif) to search for. Extracting the same features as for the search space motifs and computing a similarity measure (SEC. 5.3.1.2) to each of the entries in the database allows us to provide the user with a ranked list of results.

5.3.1.1 Feature Extraction

We found that motifs are very well described by the shape of their outlines. Thus, we base our retrieval mostly on the motifs' contours and, to a lesser extent, their colorization. In terms of shape, we found that the *Shape Context* FD by Belongie *et al.* [99] is an appropriate choice for this application. This local shape FD takes a point set representation of the contour as input. In contrast to *keypoints* ('significant' points in an image) which are used by most local FDs, the points used for the shape contexts are taken at (roughly) uniform intervals along the contour, giving a set of points $\mathcal{P} = \{P_0, \dots, P_{N-1}\}$, $P_i \in \mathbb{R}^2$. For each point of this set, a FV is computed, describing the distribution histogram of the point's neighbors in log-polar space, thus increasing the descriptor's sensitivity *w.r.t.* near-by points. An illustration of this idea is given in FIG. 5.2. In terms of histogram resolution, we use $b_r = 5$ radial bins for the log-radial component and $b_\theta = 12$ angular bins, as recommended by Belongie *et al.* [99].

To extract $\mathcal{P}$, we exclusively use the extreme external contour $\delta\Omega$ of the motif with a fixed number of points of $n_{\mathcal{P}} = 100$, which is an appropriate trade-off between accuracy and computational performance. To this end, we took every $\lceil |\delta\Omega|/n_{\mathcal{P}} \rceil$ -th point of the pixel-wise external contour, with $|\delta\Omega|$ as the number of contour points. If the resulting $n_{\mathcal{P}}$ is too high, randomly selected points are removed until the required number is met. Conversely, if the resulting $n_{\mathcal{P}}$ is too low, points are added iteratively by randomly selecting a point and placing a new point at a linearly interpolated position in the direction of the next neighbor.

5.3.1.2 Similarity Score

With a FV assigned to all points $\{Q_i\}_{i \in I}$ of a query shape and all points $\{T_j\}_{j \in J}$ of a certain target shape, the costs $C_{ij} = C(Q_i, T_j)$ of matching a pair of points can be estimated. Since histograms are compared, the χ^2 distance is an appropriate metric to this end. The total cost for assigning all points in the query to all points in the target with the assignment π is then given by the sum

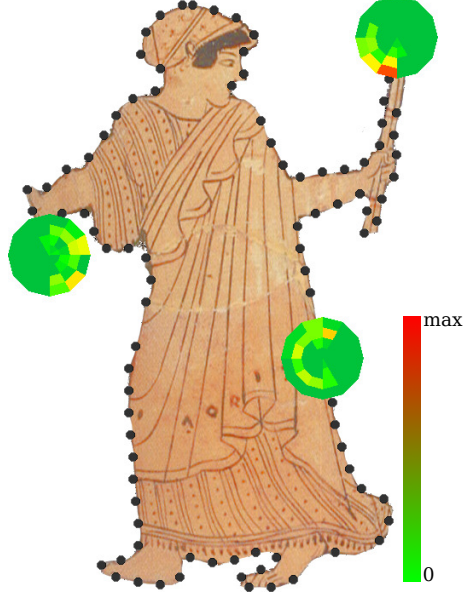


FIGURE 5.2. A motif showing a woman with a torch with its sampled contour (black dots). At three distinct locations, the point histograms of the shape context FD [99] is visualized. These polar plots in traffic light colorization indicate the direction and abundance of other points.

of the individual costs

$$(5.1) \quad \tilde{H}(\pi) = \sum_{i \in I} C(Q_i, T_{\pi(i)})$$

over all points I , which is an inverse measure of similarity. Finding the optimal assignment (the assignment with the lowest costs) is a square assignment problem. It can be solved with the Hungarian algorithm [224] with a runtime complexity of $\mathcal{O}(n^3)$. For our experiments, we applied the more efficient method proposed by Jonker and Volgenant [225], which is also of cubic complexity but drastically reduces computational load through several optimizations.

From the inspection of the search space and discussion with domain experts, we conclude that motifs often appear in similar poses but in a mirrored fashion. One specific example can be seen in FIG. 5.4, showing two men with an outstretched arm, one facing leftwards and another one facing rightwards. Hence, we make our retrieval also reflection-invariant by additionally considering a mirrored version of our query. This is achieved efficiently by reversing the order of the angular bins b_θ of the query descriptors, giving a reflection-invariant sum of costs $\tilde{H}_{ref}(\pi)$. The overall cost *w.r.t.* a specific target is thus given by

$$(5.2) \quad H = \min\{\tilde{H}, \tilde{H}_{ref}\}.$$

Sorting all motifs in our database ascendingly by H yields a similarity-ranked list of the search space. Alongside valid motif segments, our database is comprised of a number of segments corresponding to background areas in-between or around motifs, as those are not discarded

during the segmentation process (SEC. 3.2.2). Yet, those can share major parts of their contour line with a motif (complementary contour) and are thus likely to also provide a good match in terms of shape while being an undesired result. At the same time, in such cases, there is a huge difference in the colorization of the query and target segment. Thus, we combine our shape-based retrieval with a color-based criterion to discriminate complementary segments. Specifically, we compute the *Bhattacharyya distance* [226] between the 256 bin grayscale histograms of query and target segments to this end. Target segments with a distance equal to or greater than 0.6 to the query segment are excluded from the result set.

5.3.1.3 Search Interface

For the query specification, we provide a customized user interface. The idea is that users can specify the query by means of an example. To this end, they have to upload an image of a reference object featuring the motif to search for. As the depiction can contain multiple motifs, we provide the users with highlighting tools (a selection brush with adjustable brush size (FIG. 5.3, left)), allowing them to select regions belonging to the motif as well as regions containing irrelevant scenery. The selections serve as positive and negative seeds for the FG/BG segmentation by Rother *et al.* [164], which is executed in the background every time the user finishes a brush stroke. In most cases, a very coarse selection with a few strokes produces sufficiently good results. The interface shows the preliminary results of this step in a separate view (FIG. 5.3, middle). In cases where the proposed segmentation does not match the user’s expectations, the selection can be refined iteratively by adding additional brush strokes. In the example FIG. 5.3, it can be seen that the hair of the depicted character has approximately the same saturation as the remaining vessel surface, leading to an erroneous segmentation with the initial coarse selection. By explicitly marking those areas as part of the foreground mask, an improved selection can be obtained (FIG. 5.3, right). The outline of the foreground of this segmentation is taken as the query for the subsequent retrieval. Alternative to this input interface, motifs depictions that have been extracted by other means, *e.g.*, graphics editors, could be loaded and used for retrieval as well.

5.3.2 Experimental Evaluation

As a database, we select two recently published CVA fascicules with a focus on red-figured vases, exhibiting a representative range of various motifs on different vessel shapes, such as lekythoi or open shapes like kraters, cups, plates, or pyxides. Those publications are CVA Berlin 13 [137] and CVA Dresden 3 [227] (*c.f.* TAB. 2.2). Almost all the depicted vessels exhibit motifs as part of their vase paintings. From all extracted images, we discard those showing a sherd or only a fraction of an object but kept only those showing the vessels as a whole, resulting in a data basis of 57 and 42 images from CVA Berlin 13 and CVA Dresden 3, respectively. A total of 152 individual image segments are extracted using morphological transformations and 785 using the EGBIS

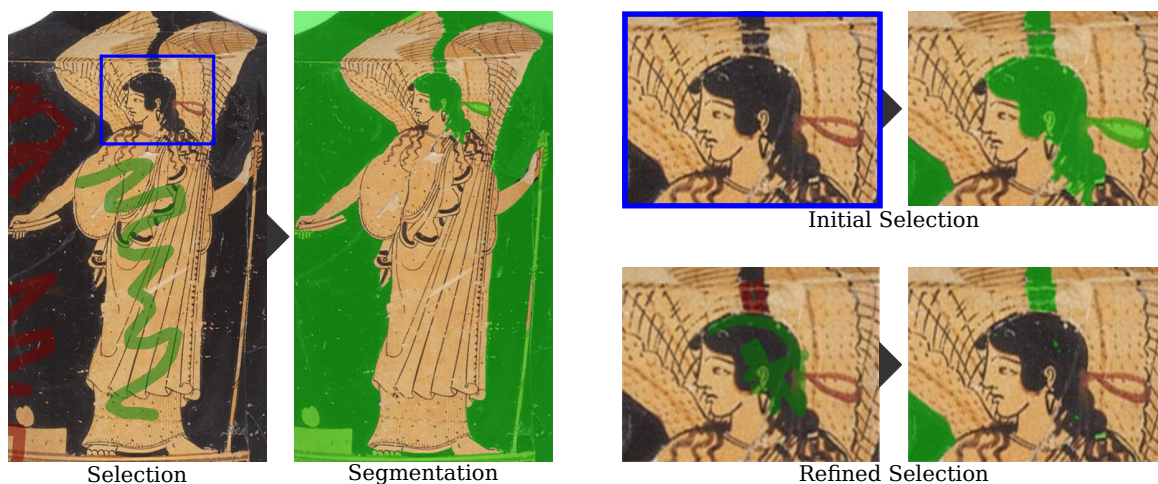


FIGURE 5.3. Left: The query motif can be specified with a selection brush of adjustable size, while the current segmentation result is displayed interactively in a separate view. The green and red brushes are used for marking relevant and irrelevant areas respectively. Generally, a coarse selection should suffice. Right: Erroneous segmentations in some areas can be iteratively refined with additional strokes.

(SEC. 3.2.2). This discrepancy results from a higher degree of fragmentation and the numerous background segments yielded by the latter. We evaluate the aspects of similarity ranking as well as different segmentation methods of our retrieval system using both figural and ornamental query motifs.

Figural Query Motifs. As a first query example, we select a depiction of a person exhibiting a characteristic pose that often appears in our domain image space. FIG. 5.4, left, illustrates the query image of a human figure with an outstretched arm to the left, a gesture of speech. The right side shows the 10 top-ranked retrieval results in descending order with the respective highlighted with a green boundary. The numbers below the figures denote the similarity score, *i.e.*, the assignment costs, EQ. (5.1)-(5.2), as well as the Bhattacharyya distance measuring the colorimetric similarity of the segments. The upper row shows the results using the baseline distance $\tilde{H}$ (EQ. (5.1)), without reflection-invariance. As expected, the top-ranked result depicts the query object itself. Overall, the top 10 results all correspond to figural motifs, with several figures exhibiting variations relevant to the query motif, as confirmed by the domain experts. Due to the reflection-variant nature of the baseline *Shape Contexts*, the found objects mostly depict figures outreaching their arm in the same direction.

Reflection Invariance. When extending the similarity score to the reflection invariant form in EQ. (5.2), we obtain the ranked result shown in the 2nd row of FIG. 5.4. A quantitative improvement of the result set can be immediately read from the tighter distribution of top similarity scores over a lower range. Qualitatively this is due to new high-ranked objects (highlighted in

blue) that particularly depict figures with an arm outreaching to the *right*. Note that, even with the exact same input query, similarity scores can vary slightly from one run to the next due to the non-deterministic sampling scheme employed for the feature extraction (SEC. 5.3.1.1). Hence, the distances of result items, appearing in both rows of FIG. 5.4 are not perfectly equal.

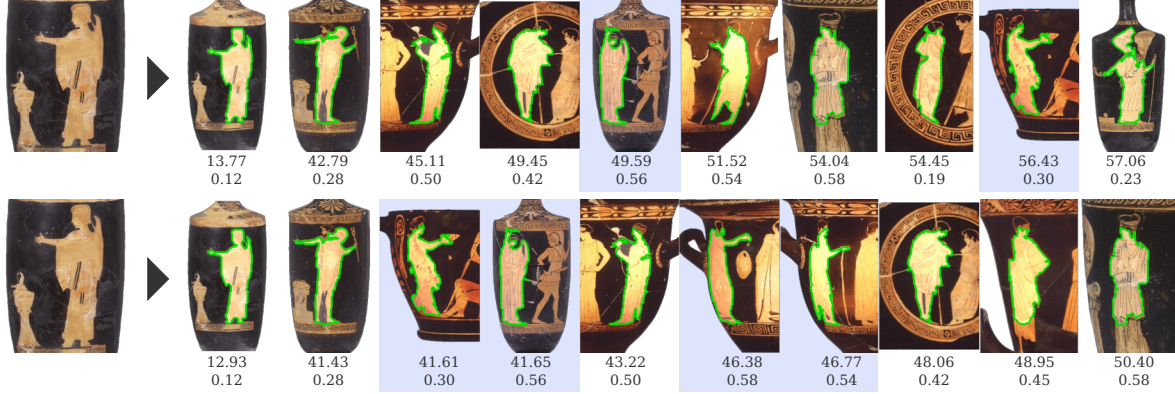


FIGURE 5.4. Query motif of a person with an outstretched arm with retrieval results and distance metrics for the baseline shape context descriptor and Bhattacharyya distance (top row) and with adjustment for reflection invariance (bottom row). Our approach is able to identify relevant motifs from the target database, including reflected motifs. Blue highlights indicate figures stretching their arm in the opposite direction of the query figure.

Another example is given by an alternative query, a palmette leaf (FIG. 5.5). Even though our method is specifically tailored to motifs as defined in SEC. 5.1, this ornament takes up a very prominent role in the vase painting while exhibiting no self-intersections. Consequently, it can be correctly segmented and searched for with our motif-search approach. This particular ornament type appears in two photographs in the search space covering two palmettes on opposite sites of the same vessel, below its handles. In this specific case, the query finds multiple similarities on both sides of the vessel, and – due to our reflection invariant descriptor – mirror-symmetric correspondences within each palmette. This example especially demonstrates the capability of our similarity search for other applications, such as the unsupervised semantics-based clustering of detail images.



FIGURE 5.5. Retrieval results for a palmette leaf appearing on opposite sides of the same vessel. Glyphs above the first seven result items indicate the front or back view of the depicted vessel.

Evaluation of Segmentation Methods. As a particularly challenging query example, we chose an often reoccurring mythological figure on Greek pottery, the depiction of Eros, the Greek god of love, appearing as a young male figure with wings. This motif exhibits a more complex shape silhouette which is particularly difficult to correctly capture, both by shape descriptors and segmentation methods. We use this example for a comparative evaluation of the two segmentation approaches we consider in our approach (SEC. 3.2.2). The first row in FIG. 5.6 shows the results for this motif using the EGBIS technique. After the query motif itself on rank one, this segmentation leads to a result set containing three further Erotes in the top 10 rankings (blue highlights). Within this top 10, we observe results depicting complex motifs that do not directly resemble a characteristic winged figure, which cannot be sufficiently discriminated by our shape descriptor. Moreover, we also find three clearly unrelated image segments corresponding to the mouths and the feet of different vases, exhibiting a similar shape as well. This is a result of our design choice to include *all* found segments in the search space. In the second row of FIG. 5.6, the same query is run on a morphologically segmented database. Compared to the EGBIS-based result, we now retrieve two Erotes, two depictions of a Nike (the winged goddess of victory), and a depiction of a Siren (a mythological hybrid of woman and bird) after the top-ranked query object. On rank 9 is an Eros image exhibiting a high similarity score that has not been present in the previous top 10 result set. As discussed in SEC. 3.2.2, this is due to an over-segmenting of its wings, which changes the overall shape. However, the shape descriptor finds hits 3 and 10 even in the face of partially degenerated segmentations. As before, we find seemingly similar motifs of complex shapes, but no unrelated mouth segment within the top 10. Overall, the morphology-based segmentation proves to perform better for this query, which is also reflected in the tighter similarity distance distribution within the top ten list.

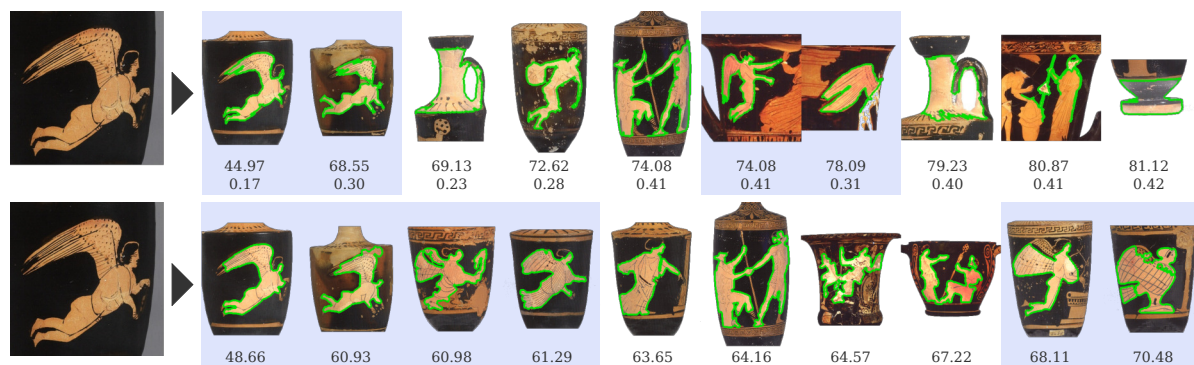


FIGURE 5.6. The retrieval results for a depiction of Eros as the query over a search space segmented with the EGBIS technique (top row) and a search space segmented with morphological operations. A blue highlight indicates the presence of a motif depicting a winged figure.

Finally, we present three additional results for queries where only one or two valid results exist in our database. Those include the motif of a Nike holding a mirror on a lekythos image

(FIG. 5.7, left). The only other Nike with a similar pose and also holding a mirror takes the top spot in the ranked results (excluding the query motif itself which is also taken from the search space). Equally good results are achieved with the second motif, the depiction of a woman’s head wearing a sakkos, a characteristic head scarf (FIG. 5.7, middle). In the third example, a man with an outstretched arm, two reference motifs appearing on the same vessel are ranked 2 and 10 (FIG. 5.7, right), emphasizing the capabilities of the reflection invariance.



FIGURE 5.7. Three additional retrieval results are presented for the query motifs depicting a Nike holding a mirror (left), a woman’s head wearing a sakkos (middle), and a man with an outstretched arm (right). To the right of the query is a selected valid result item with the number(s) in the lower left corner indicating how it was ranked.

5.3.3 Limitations and Discussion

Motifs rarely occur unattached to the surrounding scenery but are oftentimes either connected to other motifs or to other (more abstract) decorative paintings, *e.g.*, a depiction of Menelaos leading Helen home from Troy in FIG. 5.8(a), left. This poses a severe problem for the applied segmentation approaches, which do not recognize them as disjoint components. With ancient pottery, a common issue is that the surfaces exhibit characteristic degenerations in the form of cracks and chippings (FIG. 5.8(a), right). This means parts of a motif are possibly missing, and at the same time, artificial gradients are introduced, which is again a difficulty for the segmentation. Generally, the motif segmentation (SEC. 3.2.2) is the part of our pipeline with the highest potential for improvement. Both our used methods have their respective strengths and weaknesses. Yet, neither of them is capable of overcoming the mentioned limitations. Further research should include the investigation of other segmentation approaches which do not exhibit these shortcomings. The presence and locations of cracks could also be dealt with in a preceding crack detection [228] step, leveraging both colorimetric and geometric cues.

The applied shape context FD [99] is variant *w.r.t.* non-rigid transformations. Hence, similar figures but in different poses are usually not detected. For our use cases, this was not much of an issue, but if this should also be supported, the used shape context descriptor could be readily exchanged for another descriptor [229], invariant *w.r.t.* this kind of transformation. High-frequency details of a query cannot be represented appropriately since the number of contour



FIGURE 5.8. (a): Two motifs are intersecting each other (green circle) and parts of other decorative paintings (blue circle). Intersections with cracks (red circles) and other surface deficiencies are also frequent. (b): Due to the motif (woman with sakkos and torch) spanning around a substantial width of the solid of revolution, it cannot be captured with a single image, but it is cut up across several photographs.

points n_p needs to be sufficiently low for an efficient retrieval. The bottleneck in terms of the runtime is the high complexity of the square assignment. While for $n_p = 100$, feature extraction from the segmented contours takes about 7.5 ms per motif, and square assignment about 25.1 ms, $n_p = 150$ already requires 16.6 ms and 94.7 ms, respectively.

The vase paintings depicted in the photographs we use as input are distorted due to the curved surface of the vessels. On the one hand, this introduces unwanted non-linear contortions to the motifs. On the other hand, motifs usually are not well captured with a single photograph, but multiple photographs from different views are necessary (FIG. 5.8(b)). At the time the presented experiments were conducted, the affiliation of photographs to objects was not yet introduced to the used search space. It is thus possible (however unlikely) to get multiple matches for the same motif on the same vessel but from different perspectives in the retrieval result set. A more pressing issue is the distortion in some motifs, induced by the curved surface they are painted on, as well as the parts of the motif altogether missing (FIG. 5.8(b)). All of these issues can be potentially overcome by applying the *Direct Elastic Unrolling* method by Houska *et al.* [78] to all records in the search space. Unfortunately, this method was only developed after the presented feasibility study, partly as a response to the mentioned shortcomings. If the motif-driven retrieval method is further improved, it would be interesting to see how this image unrolling method can benefit the retrieval. Alternatively, if 3D models of search space objects are available, they could be included without further adaptations of the pipeline if they are flattened with any unwrapping method (*c.f.* SEC. 3.1).

5.4 Conclusion

We can summarize that – in contrast to shape-based retrieval – the retrieval based on motif similarity has not yet been the focus of intense research. Existing approaches rather focus on

classification instead of searching. To this end, they employ deep learning techniques which they train on large sets of annotated data. While this seems to result in decent precisions, the field of possible applications is very narrow as the employed networks are trained for very specific tasks. As opposed to this, our engineered approach is readily applicable for other cultures and epochs, or even other domains altogether, if the similarity of surface areas correlates with their shape and color. Since our approach is based on semantic segmentation of vase paintings, we also have to cope with the limitations and problems inherent to this process. To this end, we combine a semantic segmentation of vessel surfaces depicted on archaeological photographs with a feature-based retrieval. For the retrieval, off-the-shelf shape and color FDs are adjusted to our needs. Our experiments show that the resulting retrieval system finds semantically relevant images for different query motifs on a representative target search space. Hence, we believe that our proposed motif retrieval pipeline poses a valuable contribution to this field of research. Also, it is likely that retrieval by motif similarity will gain increased relevance due to the need by domain researchers, who face an ever-increasing number of documented artifacts, and a lack of robust methods to this end.

ORNAMENT-RETRIEVAL

Contents

6.1	Introduction	98
6.2	Related Work	101
6.3	Repetitive Pattern Recognition Benchmark	104
6.3.1	Object Selection	106
6.3.2	Annotation Concept	106
6.3.3	Annotation Tools	108
6.3.4	Evaluation Metrics	110
6.3.5	Experimental Evaluation	113
6.3.6	Possible Extensions and Outlook	114
6.4	Interactive Annotation and Classification of Geometric Ornaments	116
6.4.1	Interactive Annotation	118
6.4.2	Results	124
6.4.3	Discussion, Limitations and Open Research Questions	126
6.4.4	Applications	128
6.5	Semi-automated Creation of archaeological Drawings	130
6.5.1	Concept	130
6.5.2	Self-similarity Detection	133
6.5.3	Experimental Results	135
6.5.4	Limitations and Outlook	136
6.6	Conclusion	138

This chapter investigates the detection, recognition, retrieval, and classification of the ornamental parts of vase paintings. Literature provides no explicit definition of ‘ornament’, which is used interchangeably with ‘pattern’. The methods presented in the following are specifically for ornaments as defined in SEC. 5.1 and have been intensively researched in the second project half.

This chapter is based on:

- Lengauer, S., Sipiran, I., Preiner, R., Schreck, T., Bustos, B., “A Benchmark Dataset for Repetitive Pattern Recognition on Textured 3D Surfaces”, *Computer Graphics Forum*, 2021. DOI: [10.1111/cgf.14352](https://doi.org/10.1111/cgf.14352)
- Lengauer, S., Houska, P., Preiner, R., Trinkl, E., Karl, S., Sipiran, I., Bustos, B., Schreck, T., “Interactive annotation of geometric ornamentation on painted pottery assisted by deep learning”, *it - Information Technology*, 2022. DOI: [doi:10.1515/itit-2022-0007](https://doi.org/10.1515/itit-2022-0007)
- Lengauer, S., Komar, A., Karl, S., Trinkl, E., Sipiran, I., Schreck, T., Preiner, R., “Semi-automated Annotation of Repetitive Ornaments on 3D Painted Pottery Surfaces”, in *Eurographics Workshop on Graphics and Cultural Heritage*, The Eurographics Association, 2020, pp. 1–4. DOI: [10.2312/gch.20201285](https://doi.org/10.2312/gch.20201285)

6.1 Introduction

As opposed to motif information (SEC. 5.1) which is mostly found on Attic and Corinthian pottery, ornamental decorations appear almost universally across all cultures and epochs. Even though the preferences for specific artistic styles and shapes vary greatly, *e.g.*, geometric shapes during the aptly named Geometric period, more organic patterns during the Attic period, and a variety of different styles in Peruvian pottery, they all have in common their relative simplicity (in contrast to motifs) as well as their characteristic repetitive occurrence. *I.e.*, ornaments typically contain distinctive geometric features and often appear in ornamental bands or friezes, thus exhibiting a significant amount of symmetry and structure. At the same time, ornaments of common type can appear in varying sizes, orientations, and irregular placements, posing a particular challenge for the detection of similarities.

That is, smart tools for supporting the documentation, the analysis, and the study of ornamental information can be generalized to a large extent. Relevant tasks for archaeological research can be split into the following three categories:

T0 Detection. The automatic identification and segmentation of repetitive patterns is a particularly relevant task, serving different applications such as documentation, analysis, and retrieval. At the same time, it has a high potential for automation since it usually entails modest but tedious labor, which can be circumvented using established self-similarity detection techniques.

T1 Similarity Search. The motivation for ornament-based retrieval methods is identical to that for shape-based or motif-based retrieval methods. *I.e.*, such systems allow the discovery of similarity clusters or unveil object affiliations and can thus assist the attribution of artifacts to painters, workshops, stylistic eras, *etc.*

T2 (Statistical) Analysis. The presence of multiple instances of a specific pattern type on an object surface or also across multiple object surfaces allows for conducting qualitative analyses regarding variability and variety [74]. Said can provide cues for generating hypotheses regarding the tools used and how a vase was painted.

T3 Documentation. Similar to motifs, ornaments are part of the archaeological documentation of vase paintings. However, due to the tediousness and time involved in tracing a large number of repetitive shapes, they are oftentimes omitted from this type of documentation.

Contributions. Our contributions to the field are primarily focused on task **T1** and **T3**, while **T2** was investigated in a recent Master’s thesis [230] based on our work. That is, we propose three different concepts which address task **T1** and **T3** to different extents.

In the first study, we address the challenges of similarity detection in ornaments. Specifically, we focus on patterns that exhibit a degree of repetitiveness, *i.e.*, such which occur more than once on a certain vessel. A key prerequisite developing and evaluating new detection techniques is the availability of an expressive ground truth dataset. Since existing benchmarks did not meet our requirements – they are, for example, mostly generated synthetically and do thus not exhibit the impairments of real-world objects or comprise geometric instead of colorimetric patterns – we present a benchmark on our own. That is, the main contribution of this study is an annotated dataset comprising 82 different 3D models of painted ancient Peruvian vessels from the MAJRC, exhibiting different levels of repetitiveness in their surface patterns. To serve the evaluation of self-similarity detection techniques, our dataset was labeled by archaeologists who identified clearly distinguishable pattern classes. For those, we manually annotated all of their respective occurrences on the corresponding 3D mesh surfaces. Along with the annotated data, we introduce an evaluation metric that can rank different techniques based on the mean average precision of correctly segmented 3D mesh patches. An evaluation of different incrementally more sophisticated sampling-based detection approaches, as well as a domain-specific technique, demonstrates the applicability of our benchmark. Alongside the source code of the evaluation scripts, we provide our annotation tools for the public to extend the benchmark and further increase its variety.

In a later work we take a look at how smart tools can support the classification of patterns. To this end, we use a subset of ancient Greek art, referred to as Geometric (SEC. 1.1). Distinctive geometric patterns are typical for specific periods, regions, workshops as well as painters and are an important cue for the archaeological tasks of dating and attribution. As mentioned, the primitives of an artifact’s ornamentation can be generally classified into a set of distinguishable pattern classes, which also appear in a similar fashion on other objects. Although a taxonomy of known pattern classes is available for Geometric pottery in subject-specific publications [17], the automatic detection and classification of such patterns on archaeological photographs poses a non-trivial challenge. In the future, such a classification functionality could be provided using

specifically designed and trained CNNs. As for now, the training of such a network is not possible due to a lack of labeled training data, which at this point, does not exist for this domain context. To address this issue, we propose an effective annotation system that allows a domain expert to segment and label parts of digitized vessel surfaces interactively. These user inputs are constantly fed back to a CNN, enabling the prediction of pattern classes for a given surface area with ever-increasing precision. Our work paves the way for fully automatic classification and analysis of large object collections, which – with the help of suitable visual analysis techniques – can answer research questions pertaining to pattern variability or change over time. While the capability of our proposed annotation pipeline is demonstrated with the example of two characteristic Greek pottery artifacts from the Geometric period, the proposed methods can be readily adopted for the patternation in any other chronological period as well as for stamped motifs.

In the third and last work, we experiment with automated support for the generation of archaeological drawings. The creation of such drawings (SEC. 2.2.3) is a common practice to extract valuable information about the scenes represented in vase paintings. Generally, this task involves extensive manual labor [74], which can be supported by the automatic detection of repetitive patterns. That is, we propose a tool supporting the creation of archaeological drawings with a semi-automatic extraction of ornamental surface sections based on a combination of user-defined queries and self-similarity detection. Appropriate heuristics allow to detect the presence and positions of ornamental bands, a frequently occurring arrangement where ornamental primitives are evenly spaced along the tangential direction of a vessel’s solid of revolution. Our interactive tool allows domain experts to efficiently select ornamental queries and assess the quality of resulting similarity detections. First experiments with real-world artifacts from ancient Greek and Peruvian cultures confirm the feasibility of the approach.

Structure of the Chapter. In the following (SEC. 6.2), we discuss relevant related work pertaining to the topics of ornament retrieval and online learning. The core of the chapter is the three technical sections corresponding to our different paper contributions. In “Repetitive Pattern Recognition Benchmark” (SEC. 6.3, corresponding to [82]), we present an annotated dataset of colorimetric patterns together with a selection of evaluation metrics for the comparison of pattern detection approaches and a description of the workflow used to obtain said annotations. The section on “Interactive Annotation and Classification of Geometric Ornaments” (SEC. 6.4, corresponding to [83]) presents a concept for the ML-supported annotation and classification of colorimetric ornaments by the example of Greek Geometric pottery. The third technical section on “Semi-automated Creation of archaeological Drawings” (SEC. 6.5, corresponding to [81]) investigates how an engineered self-similarity detection together with user interaction can support the generation of archaeological drawings. SEC. 6.6 concludes the chapter with a summarization of the presented contributions and a discussion of open research objectives.

6.2 Related Work

Considering the varied approaches presented within this chapter, we focus on different scientific topics. First, the very broad topic of pattern recognition. Regarding this field, we limit the related work to CH applications, which is still an encompassing subject. We can differ between approaches for the detection of patterns on 2D sources (*e.g.*, ancient texts) and patterns on 3D shapes. The latter can be further split. While several methods have been proposed for the recognition of geometric ornaments/reliefs [231–237] (*c.f.* disambiguation in SEC. 5.1), the recognition of colorimetric ornaments [238–240] appears to be a more elusive topic. Some approaches even combine those types [233]. Second, we take a look at different benchmark datasets for the evaluation of pattern recognition approaches. Existing work on that is limited to entries to the SHREC. Finally, we discuss existing works regarding the topic of interactive ML and labeling.

Pattern Recognition. Zhou *et al.* [231] present a web-based system for matching pottery sherds, exhibiting stamped reliefs on their surface, to the respective paddle stamp design that was most likely originally used for making these specific imprints. For each 3D-scanned sherd uploaded by a user, a CNN extracts a binary image that encodes the (partial) curve patterns of the stamp design. Another CNN consumes these binary images and determines the best matching reference stamp design stored in a remote database. Curve-pattern extraction and sherd-to-design matching are performed remotely on a GPU cluster. This allows archaeologists to use the service directly from their browsers for remote collaboration.

Moscoso Thompson *et al.* [232, 233, 238] show that variants of the Local Binary Pattern (LBP), a descriptor originally introduced to describe the local distribution of intensities of a binary image, are suitable for detecting both geometric and colorimetric patterns. They extend an existing LBP variant for triangle meshes (*meshLBP*) [241] to work for arbitrary surface tessellations. This so-called *edgeLBP* achieves state-of-the-art performance for the retrieval of geometric [232] and colorimetric [238] patterns. Later on, they proposed another extension of the LBP descriptor, the so-called Mean Point LBP (mpLBP) [233]. This descriptor was designed for the retrieval and classification of objects with similar geometrical patterns on the surface of a 3D shape. The definition of pattern in this work refers to elements that are repetitive. Thus, it does not characterize local decorative patterns but the whole object. Also, the considered patterns could be geometric (*e.g.*, variations on the shape) or colorimetric (*e.g.*, paintings). The paper experimentally shows that the mpLBP is effective for retrieving relief patterns and that it is computationally less expensive than related state-of-the-art techniques like the *edgeLBP*.

Pawlowicz and Downum [239] show how CNNs allow classifying pottery sherds exhibiting colorimetric patterns based on colored images. They test their proposed system on Tusayan White Ware from Northeast Arizona and conclude that their appropriately trained classifier is able to label inputs with an accuracy comparable to a trained human.

Bogacz and Mara [234] present GigaMesh, a software framework for visualizing 3D-scanned

cuneiform tablets. The software uses the so-called Multi-Scale Integral Invariant filtering for describing the 3D data of the scanned tablets. From this input, it produces visualizations of the documents that improve their readability. Furthermore, it supports finding repetitive cuneiform patterns in documents, applying ML methods for the feature computation. The authors state that the GigaMesh framework could be extended for related problems, like studying stone wall carvings or analyzing patterns in old documents.

Úbeda *et al.* [240] propose a deep learning approach for spotting patterns in historical documents. Their approach works in two stages. The first stage indexes the documents from the collection, computing deep FVs of local regions at multiple scales using the RetinaNet network. The second stage processes and locates the patterns on the documents relevant to a query image (QBE) by similarity search. The authors perform an experimental evaluation using a dataset of hand-written medieval documents, showing that their approach obtains better effectiveness than the state-of-the-art for the pattern localization task.

In Rodriguez and Song [235], saliency features are used to retrieve similar 3D ornamentations. Visualizing the features with a feature map also supports the identification and comparison of local ornament elements. Many methods leverage a self-similarity approach to search for repetitive structures in 3D surfaces. Gal and Cohen-Or [236] propose a means to detect and describe salient regions based on curvature information. This method also proposes the construction of a hashing index to accelerate the search for similar geometric patterns. Similarly, Mitra *et al.* [237] devise an intrinsic approach to detect repetitive geometric patterns when the surface undergoes a near-isometric transformation.

Benchmark Datasets. Several entries have been made to the SHREC. The SHREC 2017 track “Retrieval of surfaces with similar relief patterns” by Biasotti *et al.* [103] proposes a benchmark for retrieving surface patches from a 3D model with a relief pattern similar to a given example. The dataset consists of 15 classes of so-called ‘tissues’, obtained from different fabrics, which define each individual relief pattern of interest. Those relief patterns are applied to non-textured 3D models (12 models per class), producing a dataset of 180 textured models. *I.e.*, the relief patterns are obtained from real-world fabrics, while the 3D models are synthetically generated. The experimental evaluation shows that the overall retrieval precision on this benchmark is rather low, indicating the task is a difficult challenge.

The SHREC 2018 track “Recognition of geometric patterns over 3D models”, also by Biasotti *et al.* [104], proposes a benchmark for recognizing relief patterns on 3D models. To this end, the track organizers built a dataset from laser-scanned fragments of archaeological artifacts, comprising 38 triangle meshes divided into a “query dataset” (8 meshes) and a “model set” (30 meshes). The task is to detect if an object contains none, one, or several of the given geometric patterns. Although several groups registered for submitting results to this track, the task proved to be too difficult for all state-of-the-art techniques, as no submissions provided fully satisfactory results. The main difficulties reported by the participants can be summed up as: (i) dealing with real-world noisy

3D data; (ii) handling high-resolution data; and (iii) the absence of an adequate training dataset. The results of this SHREC track show that the general problem of finding patterns on actual archaeological artifacts is a very hard problem to tackle, stressing the need for new datasets that can be used for testing and training new retrieval methods.

The SHREC 2018 track “Retrieval of gray patterns depicted on 3D models” by Moscoso Thompson *et al.* [105] proposes a benchmark for retrieving objects with a given painted colorimetric pattern. For this benchmark, the dataset is synthetically generated using binary patterns. The 3D models can be fully covered by a single pattern, have a faded double pattern, or be partially covered with the pattern. The organizers define two different datasets: a Single Pattern (200 models) and a Complete Dataset (100 additional models with double textures). The goal of the task is to group the objects according to their colorimetric pattern while disregarding the objects’ geometry. From six registered groups, only three submitted results for the track, revealing the high complexity of the posed challenge. All submitted results used some kind of FVs. Even though some of the methods achieved good effectiveness scores, the task proved to be challenging.

Two years later, the SHREC 2020 track “Retrieval of digital surfaces with similar geometric reliefs” by Moscoso Thompson *et al.* [106], proposes a benchmark for retrieving surface patches with similar geometric relief. The global geometry of the 3D models is not relevant to this task. The organizers of this SHREC track synthetically generated a dataset with 11 different geometric reliefs, introduced to 20 base 3D models. The patterns are purely geometric and are applied to the whole surface of the 3D model. The task is to detect groups of objects that have the same geometric relief, regardless of the shapes of the objects. To ascertain this premise, the global shape of the 3D models is not relevant for comparing the objects, but the geometric relief was the only discriminating feature. Several groups participated in this task, using a wide variety of retrieval techniques, including classic FVs and deep learning techniques. Overall, methods based on deep learning techniques obtained the best results on this track, although some of the methods based on FVs also achieved high effectiveness scores. This topic is related to the problem of finding geometric patterns in surfaces [242], which has already been studied in the context of CH [219]. More specific approaches take advantage of the symmetric structure of regular patterns in geometric surfaces to detect self-similarities [243–246].

Interactive Machine Learning and Labeling. Involving the ‘human in the loop’ of the ML process is a valuable goal for many reasons, *e.g.*, for the understanding of data and learning rules from human decision makers, providing training data, or interactively adjusting the parameters of the ML models. Also, explainability and trust in ML methods become increasingly important as more and more decisions are automatized nowadays [247]. To date, many approaches to integrate ML and interactive interfaces have been proposed. Visual analytics approaches rely on visualization to show data, models, and parameters and allow for interaction [248]. Relevant to our work are previous approaches in interactive labeling [249, 250], which aim to provide a small number of training data items that improve the ML models, *i.e.*, classification tasks in these

works. Typically, a number of views allow the user to inspect data features and classification boundaries of the trained methods. Users can explore that data, change or add labels, and compare the effects of their labeling on the underlying models.

Positioning of our work. The topic of benchmark datasets appears to be reasonably well covered by the works of the group around Moscoso Thompson [103–106], who presented several SHREC tracks regarding this topic in the last few years. However, substantial differences *w.r.t.* our proposed benchmark can be observed. *I.e.*, the presented datasets are either focused on geometric patterns [103, 104, 106] and/or generated synthetically [103, 106]. In “Retrieval of gray patterns depicted on 3D models” [105], the only benchmark featuring colorimetric patterns, the objects exhibit exclusively binary patterns. Hence, we believe that our proposed benchmark, which features *polychromatic* patterns in a *repetitive fashion on real-world artifacts*, fills an important research gap.

The online learning approach which we employ for the interactive annotation and classification is not particularly novel as the applicability of such techniques has already been demonstrated within other applications [249, 250]. It has, however, to our knowledge, not yet been employed and evaluated for CH applications. Moreover, while the considered data can be abstract and high-dimensional if suitable visualization techniques are used [249, 250], the elements to be labeled in our application are image-based, *i.e.*, patterns on 3D objects.

The last contribution pertaining to the creation of archaeological drawings relies on established computer vision techniques for self-similarity detection but presents a novel application in the form of interactive drawing.

6.3 Repetitive Pattern Recognition Benchmark

The decorative ornaments on ancient Peruvian pottery exhibit distinctive geometric shapes of varying styles, sizes, and patterns of reoccurrence (FIG. 6.1). The automatic identification of similar texture patterns on vessel surfaces can serve different important domain applications. However, state-of-the-art methods are nowhere near achieving this objective with the necessary accuracy. The development of techniques tackling this particular problem would be supported or even catalyzed by the availability of an expressive reference dataset as well as an established evaluation method for comparing approaches.

We present an expert-annotated benchmark dataset for repetitive pattern detection algorithms to address these needs. We provide an annotated dataset of different 3D models of ancient painted Peruvian vessels exhibiting varying levels of repetitiveness in their patterns. The dataset is complemented by manual expert annotations defining a reference set for pattern classes and their similarity instances on the vessel surfaces. Furthermore, we introduce an evaluation metric that can rank different recognition techniques. With this dataset and its accompanying evaluation benchmark, we hope to induce and inspire the development of novel approaches for the



FIGURE 6.1. A selection of exhibits from our benchmark dataset. The annotated classes of surface patterns (squares) exhibit different shapes and styles, while their individual entities (colored polygons) show vastly different levels of repetitiveness and order.

detection of similar surface patterns, not only in the classic pattern recognition community but also within the field of geometry processing and shape analysis, which we believe can both provide significant contributions to this research problem in the future. Our benchmark can be utilized for the evaluation of retrieval algorithms but might also be used to test novel analysis and visual exploration techniques for surface patterns *w.r.t.* their variety, variability, and similarities across different objects. Moreover, we envision novel multi-modal shape and texture reconstruction approaches for the digital restoration of CH artifacts based on pattern similarities. In summary, the contributions of this work are:

1. A dataset of 82 textured meshes with repetitive surface patterns, annotated by a total of 2,529 reference pattern occurrences, defining 102 distinct similarity classes;
2. An annotation toolset for defining such similar pattern classes for textured 3D meshes;
3. Two orthogonal benchmark evaluation methods for ranking different recognition techniques based on a suitable precision measure; and
4. An evaluation of an incremental set of naïve sampling-based pattern detection techniques as well as a domain-related detection technique tailored to repetitive ornaments.

The annotated dataset is publicly available¹ under the *Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International* license. The site also provides the source code of the tools used for obtaining the annotations and scripts for conducting the evaluation.

The remainder of the section outlines the objects we selected for annotation (SEC. 6.3.1) before the annotation concept (SEC. 6.3.2) and used tools (SEC. 6.3.3) are presented in detail. Based on a collection of increasingly complex recognition approaches (SEC. 6.3.5), we demonstrate the meaningfulness of our proposed evaluation metrics (SEC. 6.3.4). Finally, we state possible extensions and a future outlook (SEC. 6.3.6).

¹<https://datasets.cg.v.tugraz.at/pattern-benchmark/>

6.3.1 Object Selection

The dataset is comprised of 82 hand-selected models of the much larger Peruvian dataset (SEC. 2.2.5) from the MAJRC with 963 objects. This subset is obtained by filtering out all vases featuring none or very limited surface decorations. Yet, it still contains several of the different vessel shape classes (jar, pitcher, bowl, figurine, basin, pot, plate, vase) originating from different pre-Columbian cultures. The whole selection is illustrated in FIG. 6.2. Additionally to the already given normalization², we rescale their bounding box to a unit cube.



FIGURE 6.2. A subset of pottery models from the MAJRC is selected for pattern annotation. All 82 models exhibit diverse surface decoration and different vessel shapes, such as jar, pitcher, bowl, figurine, basin, pot, plate or vase.

Note that the original physical vessels exhibit different impairments visible in the captured 3D models, such as cracks and holes in the material (FIG. 6.3(a)), but also chipped-off and worn-off surface areas (FIG. 6.3(c) and FIG. 6.3(d) respectively) that deteriorate the final texture. Additionally, some objects exhibit missing surface parts due to occlusions and other errors in the scanning process (FIG. 6.3(b)). These imperfections are characteristic of captured real-world models and add another quality level to this dataset, as it poses an additional challenge to algorithms operating on them.

6.3.2 Annotation Concept

All of the selected objects feature between one and three distinguishable similarity classes of surface patterns (henceforth referred to as *pattern archetype*), each of which has between 2 and 134 (mean $\mu_{ent.} = 24.8$) separable occurrences (henceforth referred to as *pattern entity*). The

²<http://www.ivan-sipiran.com/shrec2021.html>



FIGURE 6.3. Impairments of the data incorporate both, geometric deficiencies like missing parts due to (a) decay and (b) scanning errors, as well as textual deficiencies due to (c) chipped-off or (d) worn-off parts.

pattern entities of a pattern archetype exhibit a similar color scheme and overall shape but differ in scale, orientation, reflection, and even non-affine transformations. We require a pattern to (i) have a finite extent on an object’s surface, excluding continuous friezes reaching around the whole solid of revolution; (ii) have at least two entities; and (iii) display some degree of ‘artistic richness’, omitting very simplistic paintings like delimiters or single-stroke patterns; in order to classify as a pattern archetype. In terms of archiving, we save the information regarding the location and extent of a pattern entity as a set of face IDs of the triangle mesh the entity appears on. Hence, for pattern entities we define the data structure

```
Pattern_Entity: {
  "face_ids":      [uint32],
  "orientation":   float32,
  "scale":         float32,
  "reflection":    bool
},
```

which additionally stores the attributes orientation $o \in [0, 2\pi)$, scale $s \in (0, \infty)$, and reflection $r \in [0, 1]$ w.r.t. to one selected sample entity. Similarly, the data structure for the pattern archetype

```
Pattern_Archetype: {
  "object_id":      uint32,
  "fold_symmetry":  int32,
  "sample_entity":  Image,
  "entities":       [Pattern_Entity]
}
```

includes the information to which object the annotation belongs to (`object_id`), as well as an image of one representative entity (`sample_entity`) and a list of associated pattern entities (`entities`). The integer-valued property `fold_symmetry` stores the archetype’s rotational symmetry order n , since the shapes of many archetypes have one or more symmetry axes (e.g., the sun wheel in FIG. 6.1 left). For those which do not exhibit any symmetry characteristics, this property is set to 0. Annotations are stored on the file system using the JSON format, which has two significant advantages. First, the devised data structure can be readily adjusted and extended to incorporate additional annotation properties. Second, the JSON format is well supported with parsers for all established programming languages, allowing fellow researchers to process the annotations with

their own tools if they prefer to do so.

Our first experiments unveiled that the mesh resolution of the original models, captured with a structured-light scanner, is not high enough to support precise per-face surface labeling. Hence, the raw meshes have been subjected to one refinement step of Loop’s [251] subdivision, resulting in about 130,000 faces per 3D model. Other data cleaning steps include the removal of inner surface areas (if present), the removal of non-manifold edges together with their faces, and the removal of unreferenced vertices.

6.3.3 Annotation Tools

The process of generating annotations for 3D models is a primarily manual task requiring expert knowledge, which we support with custom-made tools (FIG. 6.4). The only required input is the textured 3D mesh for which the annotation is created, and the resulting output is a JSON file containing the annotations in the aforementioned data structure. We decide to divide the annotation process over two inter-operating tools, which has the advantage that the annotation effort can be easily split up between different groups of people. The first tool (*Sample Pattern Selection*) is responsible for determining all pattern archetypes present on the input object and outputs a set of images, each displaying one characteristic entity. The second tool (*Pattern Entity Selection*) takes these images and the 3D mesh as input and yields the annotation file.

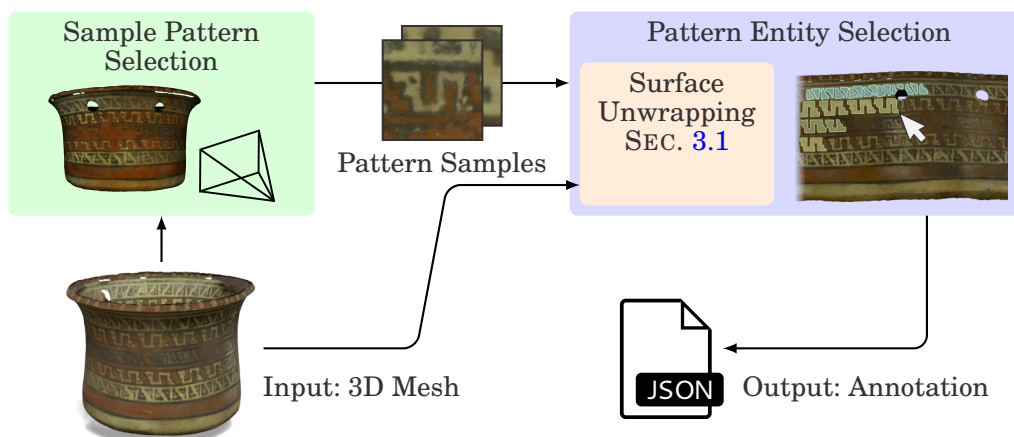


FIGURE 6.4. The annotation pipeline takes solely a textured 3D model as input from which the different pattern archetypes are identified in the *Sample Pattern Selection* step. The outcome of this step is a set of characteristic images (one per pattern archetype). This intermediate result as well as the 3D model are required for the *Pattern Entity Selection*, where individual pattern entities are identified and marked on the flattened 3D surface. Finally, the output is exported in JSON format.

Sample Pattern Selection. Let S be a 3D object’s surface. To specify a pattern archetype, we ask the user to select a set of surface points $P \in S$ around one of its entities. The first step is fitting a plane $T = \mathbf{n} \cdot \mathbf{x} + d$ to P and determining its outward-pointing normal $\mathbf{n}$ (FIG. 6.5, left).

We then project the centroid $\hat{P}$ of $\mathbf{P}$ onto S along the normal direction $\mathbf{n}$ to obtain the central selection point $\hat{S}$ on the surface. To this end, the ray $\hat{P} + \mathbf{n}t$ is intersected with S . The largest intersection parameter t_{max} then defines the surface point $\hat{S} = \hat{P} + \mathbf{n}t_{max}$.

In the second step, we set up the camera to obtain an encompassing image of the selected pattern entity that is centralized in $\hat{S}$ (FIG. 6.5, right). To this end, we create a perspective camera with 90° vertical field of view and 256×256 pixels resolution. To obtain a suitable view transformation, we first translate the shape to place $\hat{S}$ in the 3D origin and rotate it such that the normal vector $\mathbf{n}$ aligns with the $Z+$ axis. The camera origin is then offset along $\mathbf{n}$ by a distance $c \cdot m$, where $m = \max_{P \in \mathbf{P}} \|P_{xy} - \hat{P}_{xy}\|$ denotes the maximum distance of points from the view axis and $c = 1.2$ considers some padding around the pattern in the image.

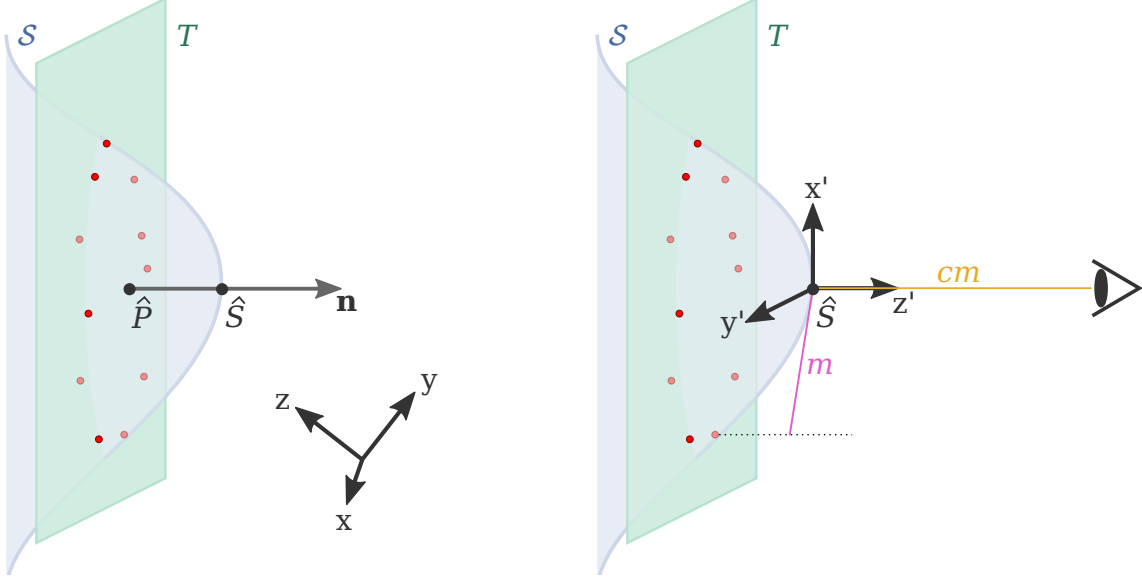


FIGURE 6.5. Left: A point $\hat{S}$ on the surface S is computed as the central surface point of the selection. Right: The camera setup for capturing the corresponding pattern selection image.

Pattern Entity Selection. For the selection of all entities of a given pattern archetype, we decided to circumnavigate the complex task of surface annotation in 3D by projecting the curved object surface to a planar proxy with our variation of the cylindrical unwrapping by Karras *et al.* [76] (SEC. 3.1). As the model’s surface needs to be ‘cut open’ at one point along the angular direction, pattern entities in the vicinity of this cut might be severed. Hence, we add additional padding to both ends of the planar surface by copying the surface, allowing for uninterrupted annotation of all pattern entities. Due to the jagged appearance of most patterns, a polygon lasso tool proved to be the most convenient tool for selecting a set of faces. All faces with at least one vertex inside the polygon are determined efficiently using a simple flood-fill algorithm.

The GUI we devise for the pattern entity selection is tripartite and includes (i) a color-coded selection list of the loaded pattern archetypes with thumbnails showing the sample images (FIG. 6.6①); (ii) a navigable graphics widget displaying the unwrapped object surface (FIG. 6.6②); and (iii) a 3D viewer, previewing the current annotation on the 3D model (FIG. 6.6③). When hovering over the unwrapped surface, the 3D model aligns such that the hovered-over position is facing the camera.

The annotation process is initiated by loading a 3D mesh, which is subsequently unwrapped, as well as a set of pattern archetypes belonging to this object, given by the sample images obtained from the *Sample Pattern Selection*. The following steps are conducted iteratively by the user until all pattern entities of all pattern archetypes are annotated:

1. (Optional) Select a pattern archetype from the selection list (FIG. 6.6①) if the next target entity does not belong to the currently selected pattern archetype.
2. Select a pattern entity on the flat proxy surface with the polygonal lasso tool (FIG. 6.6②).
3. The *Pattern Entity Selection* switches into entity edit mode, which allows specifying the orientation and scale *w.r.t.* the archetype's sample image by aligning the overlaid image with the current selection (FIG. 6.6④). At this point, it can also be declared whether or not the current entity is a reflection with respect to the archetype sample.

At the end of the annotation process, the result can be exported as a JSON file. The tool also allows to load and edit existing annotations. The appropriate selection of sample patterns is oftentimes debatable (*e.g.*, larger patterns could be split into composing parts or two types of similar pattern archetypes could be merged into a single one) and requires domain knowledge. The selection of the individual entities, on the other hand, is usually straightforward and can be conducted by laymen.

6.3.4 Evaluation Metrics

We evaluate the benchmark on a per-pattern archetype level. To this end, we assume a query, constituting either the archetype's sample image or one of its pattern entities, as input for a retrieval algorithm. Given this input, the retrieval is required to return a set of coherent surface areas $\mathbf{R} = \{\mathcal{R}_j, j \in J\}$ (detections), each of which is given by a set of face IDs $\mathcal{R}_j$ and referring to a detected pattern entity. J denotes the index set of retrieval results. Alongside the detections, some retrieval algorithms also yield associated confidence values. The first evaluation metric we propose (Mean Average Precision) follows the method presented in the PASCAL Visual Object Classes Challenge [252] for object detection tasks. For techniques that approach the retrieval problem in 2D space after a surface parametrization (*c.f.* *Pattern Entity Selection*), we provide an alternative evaluation metric (Average Nearest-Neighbor Distance) that allows evaluating the

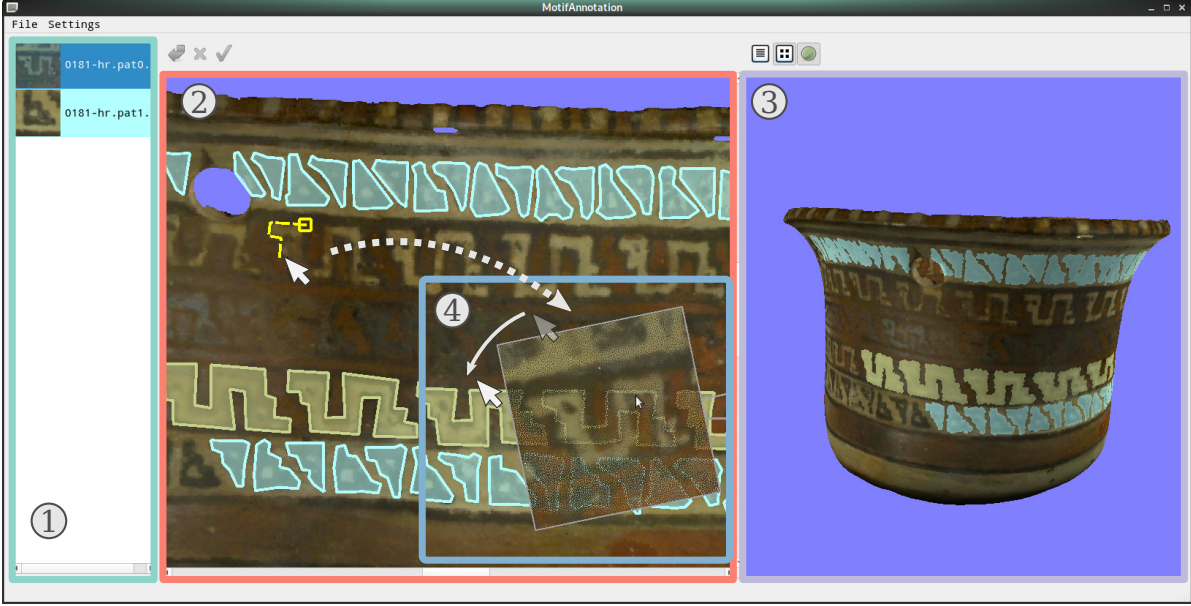


FIGURE 6.6. The *Sample Pattern Selection* with a ① colored selection list, an ② occurrence selection on the flattened surface, and a ③ 3D viewer with the annotated model. Upon finishing the polygon selection, the scale and orientation can be specified by aligning the overlaid sample pattern ④.

effectiveness of a retrieval method in image space. To this end, we assume that such a retrieval algorithm yields 2D point locations and (optionally) also affine transformations *w.r.t.* the query.

For both metrics, implementations are provided on the associated website, which can be readily employed to evaluate the performance of different retrieval algorithms. Note that the presented metrics are mere suggestions and that researchers are invited to develop their own metrics to evaluate other traits of retrieval methods and the dataset.

Mean Average Precision. Let $\mathbf{G} = \{\mathcal{G}_i\}_{i \in I}$, with index set I , denote the set of surface areas corresponding to the pattern entities of the given archetype (ground truth). We determine a correspondence between the elements of $\mathbf{R}$ and $\mathbf{G}$ and compute the Intersection over Union (IoU) rate for every possible unordered pair combination $\langle \mathcal{R}_j, \mathcal{G}_i \rangle$. Since there could be detections that have a considerable overlap with more than one ground truth annotation, we need to assign the detections to avoid multiple correspondences. We proceed greedily by first sorting the potential pairings according to their decreasing IoU rates. Thus, our method iterates over the assigned pairs while selecting one pair at a time, and the detections and patterns involved in the selection cannot be chosen again in a subsequent pair. This procedure yields a correspondence set between detections and pattern occurrences.

Next, our method sorts the correspondence set by the confidence of the detections. We also mark the pairs with IoU above a threshold t_{mAP} as true positives and detections with IoU below t_{mAP} as false positives ($t_{mAP} = 0.5$ is a good choice for practical applications and is used

throughout our experiments). The ground truths without a match in the detections become false negatives. We use the ranked list of correspondences to compute recall and precision values and subsequently calculate the 11-point interpolated precision values for the archetype pattern. Finally, the average precision for an archetype is the average of the eleven interpolated precision values. The final evaluation measure – the Mean Average Precision (mAP) – is the mean value over the average precisions of patterns in the complete collection.

Average Nearest-Neighbor Distance. The Average Nearest-Neighbor Distance (ANND) operates on the flattened object surface and returns a normalized average distance. The method can be applied for methods that return the center point locations of detections $\mathbf{C} = \{C_j\}_{j \in J} \subset \mathbb{R}^2$ and optionally also the affine transformations scale $S = \{s_j\} \subset \mathbb{R}_+$, orientation $O = \{o_j\} \subset \mathbb{R}_+$, and reflection $R = \{r_j\} \subset \mathbb{B}$, *w.r.t.* a given query. Hence, $\{C_i\}_{i \in I}$, $\{s_i\}$, $\{o_i\}$, and $\{r_i\}$ describe the location, scale, orientation, and reflection of the ground truths, respectively. The ANND for a retrieval result is defined as

$$(6.1) \quad \bar{d}_{ANND} = \begin{cases} \frac{1}{|J|} \sum_{i=0}^{|I|-1} \tilde{d}(i, \hat{\phi}(i)) + (|J| - |I|), & \text{if } |I| \leq |J| \\ \frac{1}{|I|} \sum_{j=0}^{|J|-1} \tilde{d}(\hat{\phi}^{-1}(j), j) + (|I| - |J|), & \text{otherwise} \end{cases}$$

with $\tilde{d}(i, j)$ as the normalized distance between the ground truth i and the detection j , which equals $\tilde{d}_G$, for retrieval algorithms yielding solely point locations $\mathbf{C}$. For algorithms which also provide orientation, scale, and reflection information, $\tilde{d} = (\tilde{d}_G + \tilde{\Delta}_S + \tilde{\Delta}_O + \tilde{\Delta}_R)/4$ may be used. Note that the accumulated distances of the optimal assignment $\hat{\phi} = \operatorname{argmin}_{\phi} \sum_{i=0}^{|I|-1} d_{i\phi(i)}$ of detections to ground truths is used, to prevent that multiple detections are mapped to one and the same ground truth. While $\tilde{d} \in [0, 1]$ yields the normalized distance between matched detections and ground truths, the second term $|J| - |I|$ or $|I| - |J|$ in Eq. (6.1) serves as a penalty if either too many or not sufficient detections are returned.

The geometric distance $\tilde{d}_G(i, j) = 1 - \theta(\|C_i - C_j\|, h_i)$ with the fast decaying smooth function $\theta(r, h) = e^{-r^2/(h/2)^2}$ ascertains that the distance descends towards zero for a good detection and is close to one otherwise. The respective support radii $\{h_i\}_{i \in I} \subset \mathbb{R}_+$ are given by the circumradii of the corresponding annotation polygons. $\tilde{\Delta}_S(i, j) = \max\{|s_i - s_j|, \Delta_{S_{max}}\}/\Delta_{S_{max}}$ describes the similarity of scales which is capped by $\Delta_{S_{max}} = 2$. The similarity of orientations is given by

$$(6.2) \quad \tilde{\Delta}_O(i, j) = \begin{cases} \min_{\psi \in \Psi} \{d_{\angle}(o_i + \psi, o_j)\}/\pi, & \text{if } n > 0 \\ 0, & \text{otherwise} \end{cases}$$

with n as the n -fold rotational symmetry of the pattern archetype, $d_{\angle}$ as the absolute difference between two angles and $\Psi = \{2\pi k/n \mid k \in [0..n)\}$ as the orientations of the archetype's n symmetry axes. The similarity of reflection states is defined as

$$(6.3) \quad \tilde{\Delta}_R(i, j) = \begin{cases} 0, & \text{if } r_i = r_j \\ 1, & \text{otherwise.} \end{cases}$$

6.3.5 Experimental Evaluation

Even though there are no retrieval algorithms tailored to this specific problem yet, we evaluate the dataset with (i) a Random Selection (RND) to determine a lower baseline; (ii) Randomized Local XOR (RLXOR), a very naïve approach leveraging bitwise surface comparison at randomized locations; and (iii) Repetitive Ornament Detection (ROD), our approach for repetitive ornament detection (presented for the creation of archaeological drawings in SEC. 6.5). Although these approaches operate on the flat surface image, we also remap the predicted surface areas back onto the 3D mesh in order to evaluate the mAP. For each pattern archetype, one retrieval, based on a randomly selected entity $q \in I$ as the query, is conducted.

Random Selection. With this approach, we generate pattern detections with individual locations, as well as orientations, scales, and flip-states, purely at random. Per retrieval, $N_p \sim \text{Poisson}(\mu_{ent.})$ detections are generated. The detections' locations $C \sim \mathcal{U}((0,0)^T, (w,h)^T)$ (with w as the flat surface width and h as the flat surface height), orientations $O \sim \mathcal{U}(0, 2\pi)$, and flip-states $R \sim \mathcal{U}\{0, 1\}$ follow uniform distributions, while their scales $S \sim \mathcal{N}(1, \sigma_{scale})$ follow a normal distribution with $\sigma_{scale} = 0.08$ empirically determined from the dataset.

Randomized Local XOR. RLXOR is a slightly improved version of RND. For the given query, $N_0 = 4,000$ initial detections are generated with the same distributions as for RND. For each detection j , a confidence value $\kappa_j = |\mathcal{S}_q|^{-1} \sum_{k \in [0..|\mathcal{S}_q|)} \mathcal{S}_q[k] \oplus \mathcal{S}_j[k]$ is computed with $\mathcal{S}_q$ as the set of ordered colored surface pixels which are within a radius of $r = s_j \cdot r_q$ from the location of the query C_q . $\mathcal{S}_j$ is the equal-sized set of surface pixels around the location of the detection C_j . The detections are sorted descendingly by their confidence values $\{\kappa_j\}$ before the top- $N_p \sim \text{Poisson}(\mu_{ent.})$ entries are returned.

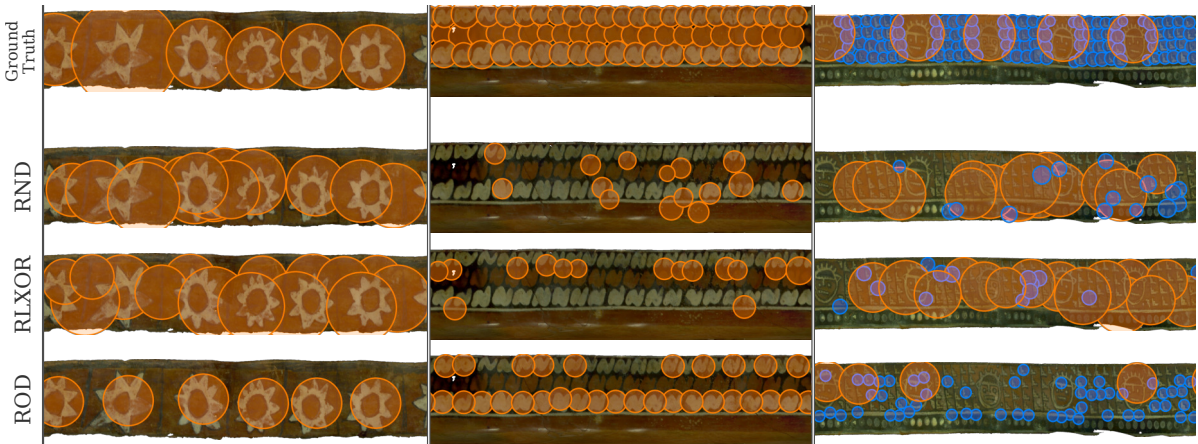


FIGURE 6.7. The comparison of retrieval results for the different query patterns given in TAB. 6.1, based on RND, RLXOR, and ROD. The colored circles mark the location and extent of a predicted pattern.

Application on the dataset. We conducted a qualitative evaluation of our metrics with three different objects from our dataset, exhibiting different quantities of patterns and different degrees of complexity. The first object displays six entities of a star-shaped pattern (FIG. 6.7, 1st column), with slightly varying numbers of jags and overall shape. While the different rows in FIG. 6.7 visualize the locations and extents of predicted pattern entities, TAB. 6.1 outlines the algorithms' quantitative performance according to the mAP and the ANND. As expected, the retrieval with RND performs very poorly according to both metrics. The RLXOR performs equally poorly, at least according to the ANND, which we attribute to the too-large number of detections. With ROD an almost perfect score was achieved with both metrics due to the relative simplicity of the vase painting and the almost bichromatic colorization.

The second object (FIG. 6.7, 2nd column) exhibits 65 very similar 'N'-shaped patterns arranged in a grid structure. As with the previous example, RND and RLXOR perform rather poorly as the number of detections differs significantly from the number of entities in the ground truth. The results obtained with ROD are approximately at the center of the scores of both metrics ([1,0] for ANND and inversely [0,1] for mAP), which meets our expectations since a third of the object's patterns (middle row) feature a rather dark hue (similar to the vessel background) and are completely lost during the necessary binarization step.

The third object (FIG. 6.7, 3rd column) poses the most challenging example, which exhibits two distinct pattern archetypes: a rather large face shape with only five entities and a smaller triangle shape with 113 entities. With this example, a slight improvement of the RLXOR over the RND is observable. With the ROD half of the entities of the first pattern are correctly detected and about a third of the second, which is clearly reflected in the evaluation scores.

Additionally to this example use cases, we perform a quantitative evaluation on the whole dataset. FIG. 6.8(a) shows the precision-recall curves for the implemented pattern detection algorithms. Generally, these results indicate that there is ample room for improvement. As expected, the trivial approach (RND) yields a very low precision for all recall levels. The second approach (RLXOR) proves to be more effective than RND, but is still far from optimal. *I.e.*, even for a recall of 0, the obtained precision is lower than 0.5 with this method, which is quite low for practical applications. The third approach (ROD) is not featured in the plot, as it was not able to yield results for many of the objects due to an improper binarization of the mostly polychromatic inputs.

6.3.6 Possible Extensions and Outlook

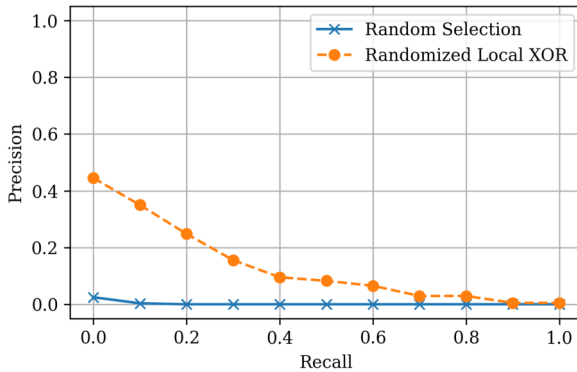
The current dataset comprises solely ancient Peruvian pottery. However, the annotation tools provided along with the dataset (SEC. 6.3.3) can be easily used to extend the collection with artifacts from other cultures or, ultimately, any kind of textured 3D meshes exhibiting repetitive surface patterns. The flexibility of our data structure and the JSON file format used for storing allow us to add additional properties relevant to other use cases. Even though patterns on Peruvian

TABLE 6.1. The quantitative performances achieved by different approaches (RND, RLXOR, and ROD) according to mAP and ANND.

								
	mAP	ANND	mAP	ANND	mAP	ANND	mAP	ANND
RND	0.02	0.71	0.05	0.86	0.23	0.75	0.01	0.92
RLXOR	0.32	0.73	0.12	0.86	0.66	0.74	0.18	0.92
ROD	1.00	0.09	0.45	0.56	0.54	0.56	0.22	0.75

pottery generally do not exhibit any identifiable hierarchical or symmetry-based relations, those could be added on top of our low-level annotation by introducing pair-wise relations between pattern entities or archetypes, such as ‘is twin of’, ‘has’, or ‘belongs to’. An example of a hierarchy is given by the three-pronged cactus shape in FIG. 6.8(b), which appears several times with three blossom shapes.

Extensions are not limited to extensions of the dataset and annotations. Note that the evaluation metrics we propose are just a starting point. For instance, they do not capture any efficiency properties of a method. Hence, arbitrary additional metrics tailored to specific requirements of a retrieval system can be defined.



(a) Precision-Recall Curve



(b) Hierarchic Compositions

FIGURE 6.8. (a): The precision-recall curves for RND and RLXOR. (b): In a few cases our assumption that pattern archetypes appear completely independent from each other (e.g., no hierarchies) does not hold, as the cactus shape, e.g., always appears together with three blossoms.

The results above show that finding repetitive patterns on vase paintings is non-trivial and requires the development of novel and robust solutions, leveraging textual as well geometric properties of the objects, for achieving high effectiveness on this task. To this end, both the dataset as well as the tools for the annotation and evaluation are publicly available. We hope that fellow researchers will come forward with innovative ideas to tackle this pattern detection problem. To date, the benchmark dataset has already found the first application within the Master’s thesis by Magdalena Mayerhofer. In the thesis titled *Visual Exploration of Repetitive Patterns in Cultural*

Heritage Objects [230], she looked into ways to visualize the multidimensional dataset enriched with well-defined intrinsic and extrinsic pattern properties. The resulting design concept is based on the clustering along two user-selected object traits. Visual links between the clusterings suggest the presence of correlations (FIG. 6.9). The prototype also supports different levels of visual granularity since a user can explore individual clusters by clicking on them or even navigate down to Single Object Preview.

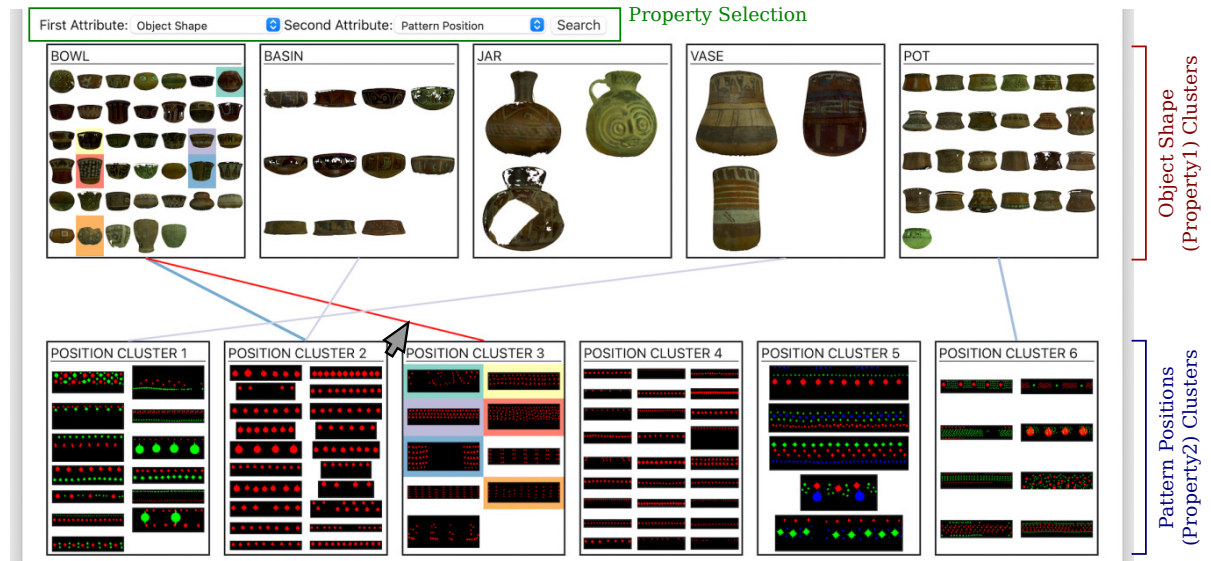


FIGURE 6.9. A collection is explored along two selected properties, which are clustered individually (top/bottom row). Visual links in-between reveal inter-cluster correlations (*Image source: [230].*)

6.4 Interactive Annotation and Classification of Geometric Ornaments

One specific group of Greek pottery is referred to as *Geometric* due to their characteristic vase paintings (SEC. 1.1). Most of the Geometric pottery is densely covered by linear ornament patterns. A categorization of these patterns helps to discover similarities and relations between objects from different sites or collections and to reveal, for example, sets of objects originating from the same region, workshop, or painter. The identification of such object relations allows domain experts to gain new insights, *e.g.*, concerning the distribution or workshop identifications.

However, to determine when and where a vase was painted, archaeologists have to resort to numerous pairwise comparisons between an object in question and plausible candidates, optimally from dated find contexts like graves or closed deposits. To qualify patterns as similar, domain experts take not only the pattern's overall appearance into account but countless other minuscule details, *e.g.*, alternating stroke directions in the hatching of a meander. When analyzing

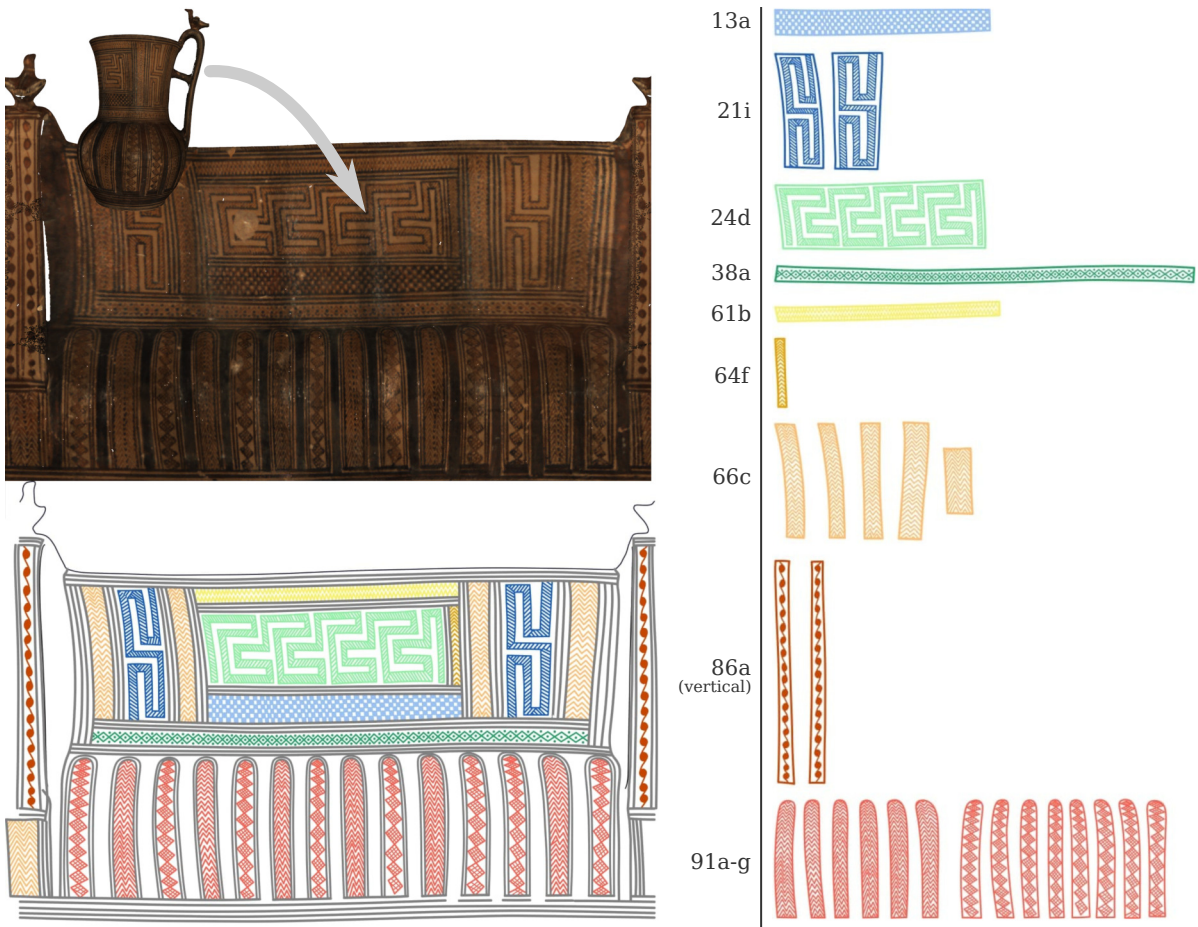


FIGURE 6.10. Left: The painted object surface after unwrapping the scanned 3D model (top) and the in-place graphical annotation of the flat surface by the archaeologist (bottom). Right: the patterns are grouped by common pattern class and classified using the Kunisch [17] taxonomy, to which the class labels refer.

a vase painting, experts typically determine (i) which distinguishable pattern classes are present; (ii) which given patterns are similar to those described by Kunisch [17] (the standard taxonomy for Geometric-era pottery, SEC. 2.1.2); (iii) which pattern classes appear together on the surface; and (iv) on which part of the vessel the patterns are located, *e.g.*, neck, body, handle, *etc.* FIG. 6.10 shows a thorough analysis of the flattened surface of the Attic Geometric pitcher *KHM IV 1*. An apparent disadvantage of this purely manual workflow is the entailing time investment, resulting in many objects ending up without proper surface documentation, leading to possibly countless undetected object relations. Moreover, in the absence of structured annotation and labeling, searching for combinations of patterns on surfaces is very complicated or altogether impossible.

We propose a structured, computer-aided annotation pipeline that requires minimal effort from the user and ensures quantitative reproducibility of annotations through well-designed

similarity metrics. To this end, a form of automatic detection is required. With established engineered features, it is often hard to formulate – and anticipate – the minuscule details required for reasonably accurate detection and classification. And often they suffer from robustness issues in the face of eroded and incomplete patterns. Conversely, ML-based techniques, CNNs in particular, have received ever-increasing attention in recent years and constitute the state-of-the-art for image detection and classification tasks [253]. Yet, same as with all other retrieval scenarios, their application is limited by the availability of appropriate training data, which currently does not exist in sufficient quantity for the domain problem relating to Greek Geometric pottery.

Hence, our idea is that: (1) an annotation pipeline in its initial state resembles the current manual analysis process; but (2) a CNN in the background learns from every action the user performs, gradually improving; and (3) eventually converging towards an expert system which can recognize pattern elements with high accuracy.

The remainder of the section is structured as follows: SEC. 6.4.1 discusses our proposed annotation pipeline in depth; SEC. 6.4.2 shows some first results, which are discussed in SEC. 6.4.3; and finally, SEC. 6.4.4 outlines further potential applications of an annotated dataset.

6.4.1 Interactive Annotation

The design of our proposed pipeline assumes that users upload a digitized version of the object they want to annotate via a custom-made annotation UI (SEC. 6.4.1.2). Our annotation system expects an image as input. Nonetheless, 3D models can also be used if they are mapped to image space (SEC. 3.1). The annotation is conducted by the alternating selection of a surface pattern and the assignment of the corresponding pattern class from a list. Upon selecting the surface area, the continuously growing database of pattern classes and annotations (SEC. 6.4.1.1) is scanned in the background, yielding a list of the most likely pattern classes based on a similarity computation (SEC. 6.4.1.4) of features obtained from a CNN (SEC. 6.4.1.3). To support interoperability and easy collaboration among experts, we decide on a web-based application that connects to a common database. For the straightforward integration of the annotations in other applications, we support the standardized export of the database.

6.4.1.1 Initial Dataset

The system supports two distinct surface annotation tasks to support the traditional manual annotation workflow, allowing us to associate surface areas with vessel parts and pattern class. To this end, the system comprises two different taxonomies for the respective annotation tasks:

Vessel Parts. Our system offers a set of vessel part labels for referring to the different morphological parts of a pottery object. The taxonomy of these parts, following the human body structure (foot, belly, neck), is widely used in pottery studies [254]. We compose an initial

label set $\mathcal{V}$, containing the most commonly used labels, beginning bottom-up: *foot*, *lower part*, *belly*, *shoulder*, *neck*, *rim*, and *handle*. However, this set is not encompassing. Hence, we allow the user to dynamically add additional labels on demand.

Geometric Patterns. Our initial set of pattern classes is based on the taxonomy by Kunisch, presented in *Ornaments of Geometric Vessels* (in German) [17]. The 776 classes, grouped into 6 major and 20 minor groups, are extracted as outlined in SEC. 2.3.1.2. Each record in this set includes (i) a representative drawing of the pattern; (ii) a textual description comprising different labels such as ‘hatched’ or ‘vertical’; and (iii) (optionally) a reference to occurrences of the pattern in literature.

6.4.1.2 Annotation Interface

For the annotation, we design a web interface where the uploaded image is displayed in the left half of the screen (FIG. 6.11 ①), while the right half of the screen shows a ranked list of the best matching pattern classes for the currently annotated area. Area selection is accomplished using a lasso drawing tool, allowing the selection of arbitrarily shaped areas of the object surface. The same lasso tool is also used for marking vessel parts. Switching between both modes is accomplished by clicking a tab right above the photo display area. The pattern classes are organized into their *minor groups*, and the best-matching pattern groups are listed from top to bottom (FIG. 6.11 ②). The assignment of a pattern class to the current selection is enabled by the respective blue checkmark buttons. If annotations for a specific pattern class have already been obtained from other inputs, those annotations are listed next to their sample pattern and sorted (from left to right) according to their similarity to the current selection (FIG. 6.11 ③). Only the best-matching candidate per group is shown by default, but the other patterns of the group can be displayed on demand (FIG. 6.11 ④), similar to the nested folder structure shown in file browsers.

6.4.1.3 Extraction of Image Features

For the similarity computation between a selection and the patterns in our database, we need to extract descriptive image features with our CNN. Pottery surface depictions are usually documented as color and grayscale photographs. Hence, we apply a grayscale conversion [102] to the color photographs to bring the inputs into a common base representation. While this simple preprocessing step suffices for the comparison among photographs, we also have to compare photographs to the archetype drawings by Kunisch for pattern classes that do not have any annotations yet. The direct feature extraction from those inputs of a different nature yields vastly different FVs. Consequently, we insert an additional preprocessing step in such cases to bring the query and reference into a common base representation (sketch representation).



FIGURE 6.11. The web-based UI comprised of an annotation area ① with an active selection and results area (right), with the best matching pattern group ②. The previously annotated inputs for a pattern class are displayed next to the pattern ③. Further pattern classes belonging to the same group (here ‘rectangle’) can be displayed on demand ④. By clicking the respective blue checkmark, the user can assign the matching pattern class for the current selection.

Sketch-like Input Abstraction. The process of transforming the archetype drawings into realistic grayscale images is an ill-posed problem. Hence, we go the other direction and mimic the artistic style used in the documented pattern taxonomy, done in a similar fashion by Lee *et al.* [79]. First, we calculate the discrete Laplacian of the grayscale image. Second, Otsu’s method [167] is applied to the inverted Laplace-filtered image, which yields a strictly binary black-and-white image. Otsu’s threshold is also applied to the Kunisch pattern images, which are semantically already binary images but may still exhibit inadvertent intensity variations due to the digitization process. FIG. 6.12 (from left to right) shows the inputs, intermediate processing steps, and the final outputs calculated by our method for a Kunisch pattern (top) and a sample annotation from a photograph (bottom).

Residual Neural Network. There are many options for selecting a CNN to extract features from images. We opt for a well-established architecture that provides high effectiveness and can be easily trained for custom purposes. The Residual Neural Network (ResNet) is an architecture that combines convolutional layers and shortcut connections, making it possible to train very deep models without compromising the convergence of the learning process due to vanishing gradients. We train several networks (ResNet18, ResNet50, ResNet101, and ResNet152, where the number at the end denotes the number of layers) and evaluate the best option for our needs. We use

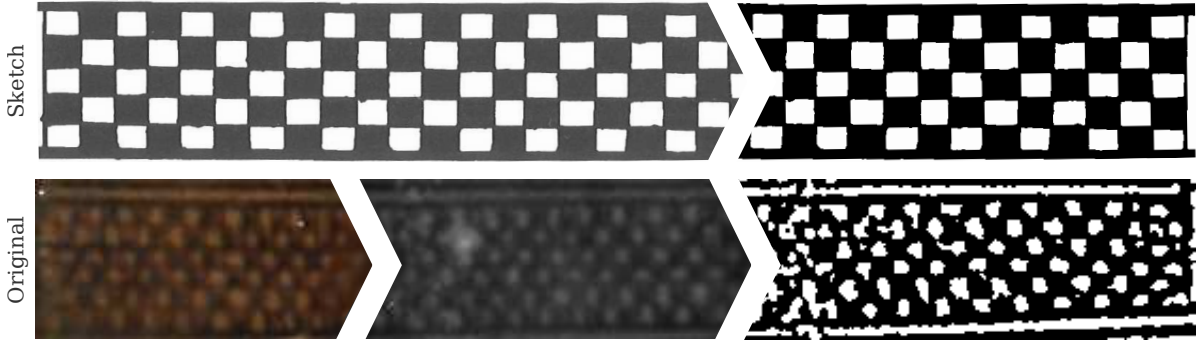


FIGURE 6.12. Kunisch pattern classes, as they are depicted in the taxonomy, are mapped from grayscale (top left) to black and white (top right). Query selections from input photographs (bottom left) are first converted to grayscale (bottom center) and finally to black-and-white (bottom right).

the Pytorch³ pre-trained networks to perform the fine-tuning. The setup for all the networks is: data augmentation (vertical and horizontal flips), batch size of 32, and Adam optimizer with a learning rate of $1e-4$.

We take the Kunisch patterns as the input dataset. To make a fair comparison of our implementations, we split the patterns into 70%/20%/10% for training, testing, and validation, respectively. Since all the networks are pre-trained with ImageNet⁴, our goal is to fine-tune them on the Kunisch patterns. ImageNet is a large-scale network trained on 1.2 million natural images. The pre-training step aims at finding the best latent representations for the images in a classification setup. Even when the network is pre-trained on this large-scale dataset, it is expected that the network has learned to extract meaningful information from the images for the downstream task. Transfer learning is a well-established methodology taking advantage of a pre-trained network to create a specialized neural network for a small, related problem. In our case, the network gains the specialization to represent the patterns via a fine-tuning of the pre-trained ResNet. Each network is trained (using the same random seed) with the training dataset for classification, while the validation set is used for fine-tuning hyper-parameters. Once the networks are trained, we drop the last classification layer from the network and use the output of the global pooling layer as FV. FIG. 6.13, left, shows the precision-recall plots for the trained networks computed over the test set. From this figure, we can conclude that 18 layers is the best option to learn from the patterns. It seems deeper networks, such as ResNet50, ResNet101, or ResNet152, require much more data to converge and to generalize, while shallower versions of ResNet would not be able to converge for learning a large-scale dataset and would thus not be useful for transfer learning. After deciding on a certain depth (18 layers), a number of runs of the same ResNet18 base network are conducted as the transfer learning uses several random data augmentations. After trying different configurations, we perform a classification

³<https://pytorch.org/>

⁴<https://www.image-net.org/>

assessment over the validation set. FIG. 6.13 also shows the 11-point precision-recall curves for other runs of ResNet18, labeled ResNet18_*, where the respective four-digit number refers to the network’s classification accuracy on the validation set. It can be seen that certain configurations entail a substantial increase in performance. It appears that ResNet18_8158 is clearly the best, which is also confirmed by established evaluation measures – Nearest Neighbor (NN), First-tier (FT), Second-tier (ST), and Mean Average Precision (mAP), details on which are given in Sipiran *et al.* [42] – shown in TAB. 6.2, where this network outperforms all others in every single metric. FIG. 6.13 right shows the confusion matrices for ResNet18_8158, computed from the nearest neighbor pattern in the retrieved list. A high NN measure ensures that the most similar pattern is retrieved in the first position in a search. It can be seen that the network performs (unsurprisingly) much better on certain pattern groups, such as rectangular or circular patterns, than on the rather complex lozenge or pictogram shapes.

Additional annotations resulting from the user-driven documentation process can be leveraged to further fine-tune the network. Since re-training the network is resource expensive, the network is re-trained only once a certain number of new samples are available. This phenomenon is known as concept drift, and there are several ways to address it [255].

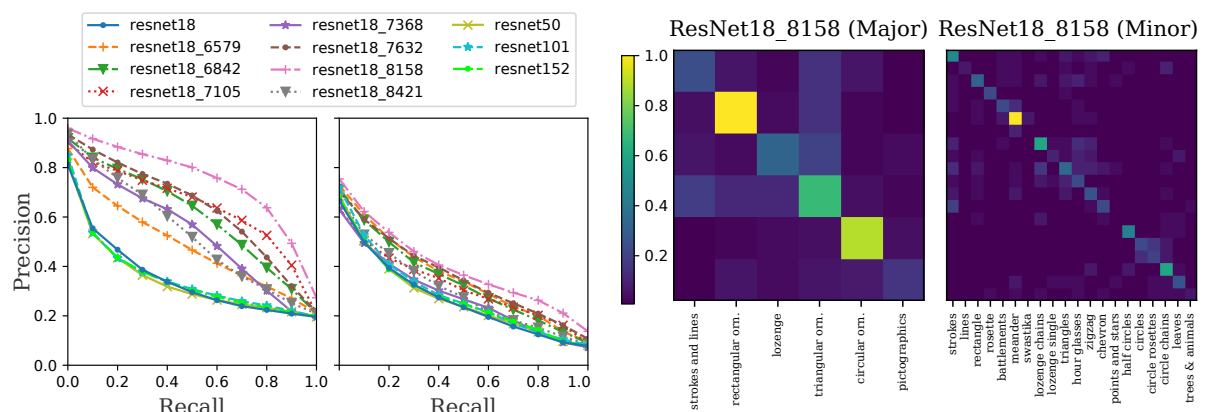


FIGURE 6.13. Left: The 11-point precision-recall curves for the major (left) and minor (right) pattern groups for different levels of depth of ResNet (18/50/101/152) and different configurations. Right: The respective confusion matrices for ResNet18_8158 and major/minor groups.

6.4.1.4 Similarity Computation

To propose accurate pattern class predictions for a surface region selected by the user, we leverage visual similarities, as well as the position on the object surface. For the prior, we employ the 512-dimensional image features $\{\mathbf{x}\} \subset \mathbb{R}^{512}$, obtained from the ResNet18_8158 network (SEC. 6.4.1.3), which we extract for the current selection. To this end, we use the polygon bounding box of the user’s lasso selection to crop the annotation surface and mask all image parts outside the selection polygon. The FV $\mathbf{x}_q$ obtained from this query is compared to all the pattern class samples

TABLE 6.2. Evaluation measures (NN, FT, ST and mAP) for all investigated ResNet architectures. The respective highest values are in bold font.

Architecture	NN	FT	ST	mAP
ResNet18	0.6695	0.3139	0.2536	0.3332
ResNet18_6579	0.7500	0.4376	0.3289	0.4709
ResNet18_6842	0.8506	0.5471	0.3600	0.5971
ResNet18_7105	0.8218	0.6002	0.3784	0.6341
ResNet18_7368	0.8305	0.4838	0.3228	0.5228
ResNet18_7632	0.8822	0.5743	0.3714	0.6297
ResNet18_8158	0.9339	0.6828	0.4127	0.7438
ResNet18_8421	0.8822	0.4688	0.3202	0.5216
ResNet50	0.6753	0.2992	0.2476	0.3247
ResNet101	0.7299	0.3113	0.2556	0.3348
ResNet152	0.6954	0.3097	0.2558	0.3321

provided by Kunisch $\{\mathbf{x}_i\}_{i \in I}$, with I as the index set of pattern classes, as well as all previously annotated patterns $\{\mathbf{x}_j\}_{j \in J_i}$, with J_i as the index set of annotations belonging to the pattern class i . We differentiate between FVs obtained from the grayscale inputs ($\mathbf{x}^G$) and FVs obtained from the binarized (SEC. 6.4.1.3) inputs ($\mathbf{x}^B$).

We found that the similarity between a query q and a pattern class i is strongly correlated with the similarity of the query to the class’s existing annotations, while a weaker correlation was observed between a query and the pattern class samples. We attribute this to the fact that they are artistic representations of a pattern and, thus, different in nature. Hence, we decide to apply a stronger weighting to the similarities of annotations if such are present. *I.e.*, we define the similarity of query q to pattern class i as the inverse of the distance

$$(6.4) \quad d(q, i) = \begin{cases} \frac{\sum_{k \in [0..|\mathcal{X}_i|)} \mathcal{X}_i[k] w(k)}{\sum_{k \in [0..|\mathcal{X}_i|)} w(k)} & \text{if } |J_i| > 0 \\ \theta(\mathbf{x}_q^B, \mathbf{x}_i^B) & \text{otherwise,} \end{cases}$$

with $\theta(\mathbf{x}_1, \mathbf{x}_2) = 512^{-\frac{1}{2}} \|\mathbf{x}_1 - \mathbf{x}_2\|$ as the normalized Euclidean distance. We apply a weighting to the distances in the sum in EQ. (6.4) relative to their appearance in the sorted sequence $\mathcal{X}_i = \text{sorted}(\{\theta(\mathbf{x}_q^G, \mathbf{x}_j^G)\}_{j \in J_i})$. The underlying idea is that the presence of a few very close matches is more important than the overall mean similarity. To this end, we employ a (fast) decaying weighting function $w(k)$, which puts increased emphasis on the best matches. $w(k) = k^{-1}$ is used in all subsequent experiments.

Often the placement of a pattern on the object’s surface provides a vital cue of the similarity computation. From the Kunisch taxonomy, we have no information regarding the positioning of certain patterns. For the annotations, however, we have this information due to the surface part annotation. Thus, for each annotation j , we can determine a placement histogram $\mathbf{s}_j \in \mathbb{R}^{|\mathcal{V}|}$, where the k -th bin $\mathbf{s}_j[k] = |\mathcal{A}_j \cap \mathcal{S}_{\mathcal{V}[k]}| / |\mathcal{A}_j|$, $k \in [0..|\mathcal{V}|]$, reflects the annotation’s correlation with the k -th surface part label. $\mathcal{A}_j$ denotes the set of surface pixel of annotation j , and $\mathcal{S}_v$ the set

of surface pixels of vessel part $v \in \mathcal{V}$. That is, the distance function θ can be reformulated to $\theta(\mathbf{x}_1, \mathbf{x}_2) = \alpha(1 - BC(\mathbf{s}_q, \mathbf{s}_j)) + (1 - \alpha)512^{-\frac{1}{2}} \|\mathbf{x}_1 - \mathbf{x}_2\|$ for the comparison of annotations. BC refers to the Bhattacharyya coefficient [226], which we employ for the distance computation between placement histograms, and the weighting factor $\alpha \in [0, 1]$ governs the influence of the placement on the similarity computation. The optimal value for α could not be determined within the scope of our experiments as it would require looking at a large number of annotations. First experiments, however, indicate that the placement should play a less substantial role, *i.e.*, $\alpha < 0.5$, than the pattern similarity.

6.4.2 Results

For the first experimental evaluation of our annotation pipeline, we investigate four different patterns of varying complexity that are found on the surface of the Attic Geometric pitcher *KHM IV 1* (FIG. 6.10). From the unassisted manual annotation effort, it is known which pattern class from the Kunisch taxonomy they can be attributed to. For the first scenario, we try to identify the correct pattern class based solely on the user selection of the respective surface area and the extracted sample pattern.

The first pattern that we look at is a right-turning hatched meander in vertical panels (according to Kunisch, no. 21i), which appears two times on the left and right side in the neck zone of the pitcher (FIG. 6.14, left). In the obtained results, the group of meanders is ranked as the closest match, while the exact class within the group is at the fifth position of the best-fitting entries. Better yet, the second best entry is actually the same pattern type, except for the turning direction (left instead of right, FIG. 6.14). The chessboard pattern in a horizontal panel (no. 13a) on the lower neck area of the vessel belongs to the group of rectangular patterns, which is ranked third out of all twenty groups by our retrieval system. Even though the correct pattern class only appears towards the end of the third quarter of the respective group, there are already several good matches among the first few classes (within the first quarter) in the group – *c.f.* the chessboard pattern at the eighth place (FIG. 6.14). For the lozenge chain with dots within and without in a horizontal panel (no. 38a) at the bottom of the neck zone, the group affiliation was correctly calculated, and the specific lozenge chain at the sides appears already within the first entries of the group. Matching the vertical tongues patterns with stacked M-chevrons and cross-hatched lozenge chains, alternately decorating the belly zone of the vessel, is especially challenging, as the provided taxonomy by Kunisch [17] has no entries that describe the displayed patterns exactly (they show other fillings). According to the domain expert, the depicted patterns correspond to no. 91a to 91g. Nonetheless, these closest equivalents (FIG. 6.10) are found within the first quarter of all the group instances.

In the second scenario, we enriched our initial database with a few real-world artifacts from external sources for several pattern classes. In order to associate the pattern-class depictions on these external vessels, we load the respective vessel photographs into our system and annotate

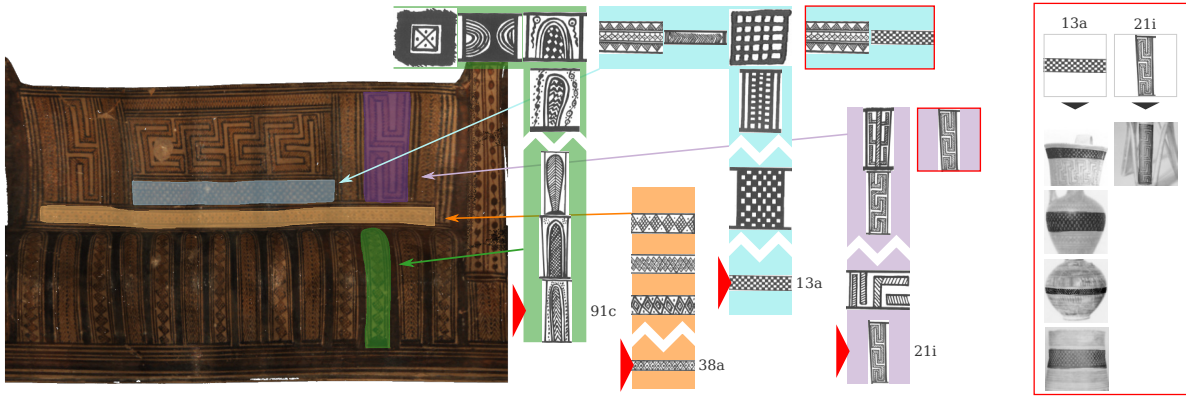


FIGURE 6.14. Surface annotation of *KHM IV 1* for four of the object's patterns. The colored panels indicate the group-wise (horizontal) and pattern-wise (vertical) ranking of the correct result according to our automatic recognition. In a second step, the same recognition was tested with a few annotations from other sources being present (red box on the right), yielding the rankings marked with a red border.

the surface areas that contain the pattern classes of interest. When we then select the same meander pattern (no. 21i) as in the first scenario, the search result improves substantially. The correct pattern class is now the top result, even after adding just a single annotation of the same meander on a different vessel (FIG. 6.14, red frames). For the chessboard pattern (no. 13a), four additional occurrences on vessels from different sources are annotated. Again, an improvement is achieved since the correct group is now listed as the second instead of the third best match, and furthermore, the associated pattern class is the best match within this group.

FIG. 6.15 illustrates the annotation process on another geometric pottery artifact, documented by Kübler [256]. This belly-handled amphora *Athens KER 2146*, decorated in three zones (belly, shoulder and neck), exhibits six clearly distinguishable classes of surface patterns (attribution to Kunisch's taxonomy done again by domain experts): (i) central on the belly a hatched meander shape in a vertical panel (no. 21h); (ii) to the left stacked cross-hatched triangle in a vertical panel (no. 47j); (iii) to the right antithetic chevrons with solid filling in a vertical panel (no. 65f); (iv) further outwards on this belly zone each one large concentric circle enclosing a reserved cross (no. 79d); and (v, vi) below these circles vertical chevrons in horizontal panels, one facing to the right (no. 65d), the other to the left (no. 65e). The whole decorated zone on the belly is framed above and below by ornament bands circumferential the whole vessel with the so-called dogtooth pattern, a row of filled triangles, in our case pointing upwards (no. 53b). Another clearly visible ornament is depicted in the neck zone: a so-called battlement pattern, on this vessel with multiple lines with dots in the open spaces (no. 19b). We annotate those patterns with our system without this domain knowledge and any prior annotations from other sources. FIG. 6.15 shows the outcome of these annotation efforts, where the colorized vertical panels to the left and right of the vessel visualize the query results for the respective surface area marked with the corresponding color. The colon-separated numbers at the top of each of the panels reflect the ranking of the matching

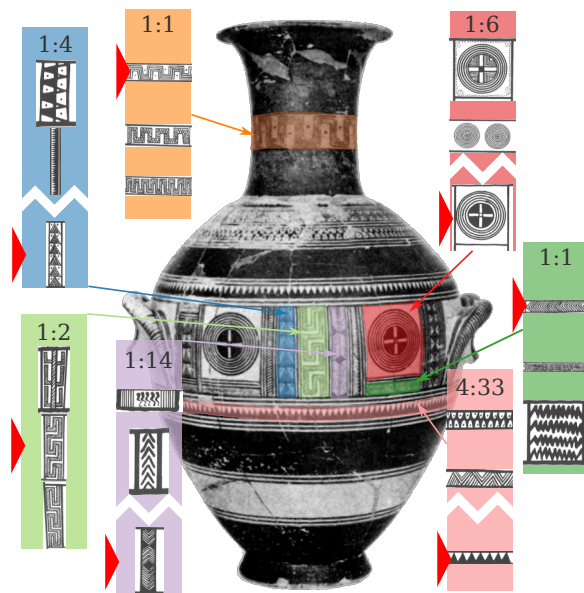


FIGURE 6.15. Surface annotation of *Athens KER 2146* [256] with the column-separated number at the top of the panels reflecting the group-wise and pattern-wise rank of the correct pattern according to our recognition.

search result for the pattern's group and the specific class within the group, respectively, and the two red arrows signify the correct class for the pattern in question. The first four of the six investigated patterns are predicted with high accuracy, as in these cases, the correct group is identified, and furthermore, the correct pattern within the group is part of the top results. The fifth pattern – the vertical chevron facing to the left in a horizontal panel (no. 65e) – provides acceptable results: while the correct group is predicted, the correct class within the group is not among the top ten candidates. The sixth pattern – the upwards-pointed dogtooth ornament (no. 53b) – is a challenging case, as our recognition system ranks the respective group of zigzag patterns only in fourth place, and the correct pattern class within the group is ranked among the classes with a low similarity measure.

6.4.3 Discussion, Limitations and Open Research Questions

From the scenarios described in SEC. 6.4.2, we can conclude that – even without any annotations – the recognition works fairly well for the majority of cases. In some cases, *e.g.*, for the dogtooth ornament (no. 53b), the correct pattern was rated as being very dissimilar, even though the preprocessed query and the reference pattern from Kunisch appear very much alike. The reason why visually similar-looking inputs lead to vastly different feature vectors on some occasions could not be established. Yet, it was found that adding annotations improved the performance in general. This is in line with our expectations, as query and search space entities are of the same modality in that case. At the same time, it was observed that a pattern with annotations

is not automatically ranked higher than all classes without any annotations, proving that the query preprocessing serves the intended purpose. Note, however, that the hand-drawn reference patterns from Kunisch are only used to address the cold-start problem [257], a well-known issue with content-based recommender systems. In the long run, we plan to replace them entirely with annotated patterns taken directly from reference photographs. This way, we will have only a single input modality throughout the entire system.

The comparison of the two types of inputs (flattened 3D model and photograph) used for the first results shows no clear preference for either input type in terms of recognition performance. However, for the practical application, the flattened surface has the advantage that it exhibits less perspective distortion and that it represents the continuous object surface as a whole.

Although the annotation was quick and smooth for the shown results, a thorough evaluation of the annotation efficiency with our system remains a future work item. *I.e.*, we intend to assess the usability of our system by conducting an encompassing user study with domain experts.

The second group of limitations is related to the algorithmic capabilities of our applied feature extraction (SEC. 6.4.1.3). The employed CNN is an interim solution due to the small number of available training data, and currently may not be most specific to our domain application. The CNN is not used as a classifier but rather only yields a latent vector for a provided input. This is reflected in a somewhat low robustness with some inputs since a slight change in the input selection occasionally results in a vastly different ranking. Additionally, inputs of a different nature (photograph and sketch) have to be compared initially, which is addressed by our preprocessing (SEC. 6.4.1.3) but still poses a challenge for certain inputs. However, contrary to the input-related limitations, these algorithmic limitations should be able to be overcome with increasing annotations, which allows the training of a more specialized CNN or even a pattern class classifier.

Automatic Pre-Segmentation of Patterns. At the moment, we require the user to manually encircle patterns to be queried. In the future, we would like to investigate user assistance in this step by proposing a set of pattern candidates that are found in the input images automatically. Even though the topic of image segmentation – the splitting of images into meaningful parts – has been a heavily researched computer vision topic for decades, specialized segmentation tasks (like in our case) still pose a significant challenge. Conventional methods comprise approaches based on thresholding, k -means clustering, and watershed algorithms [102], while modern concepts almost exclusively rely on deep learning [258]. However, of-the-self methods are not able to segment the vessel surfaces into their pattern atoms, and learning customized segmentation models requires a large amount of training data. Hence, an automatic pre-segmentation of patterns can be feasible once we have collected sufficient data. Yet, even with this potential enhancement, we believe that the user should always have the final choice, such that inadequate automatic segmentation results can always be edited or discarded altogether.

Pattern Recognition. The pattern recognition part of our pipeline faces two distinct groups of limitations. First, limitations inherent to the input data. Those can be further split into issues related to the artifacts themselves, which result from poor conservation and include phenomena like chipped or worn-off surface parts and missing parts in general, as well as limitations resulting from the digital data representation. Latter depends on the form of the input data and comprises various acquisition errors, *e.g.*, occlusion, registration errors, *etc.*, in the case of 3D digitized objects, and all the challenges characteristic for photographs, resolution, quality, perspective distortions, highlights, *etc.*, in case of images. Second, another data-related limitation specific to our application is that we initially do not have any locality information (which is a vital clue to exclude certain search results) for the pattern classes, as this information is not given in the Kunisch taxonomy.

Kunisch Taxonomy. Another group of limitations stems from basing our pattern classes on the Kunisch taxonomy. This taxonomy was seemingly compiled with a manual annotation process in mind. Hence, it lacks the logical, bottom-up hierarchy necessary for automatic classification. *E.g.*, several pairs of pattern classes only differ in their orientation (vertical or horizontal) while rotation-invariance is generally required. A more logical representation of this particular circumstance would be to have just one pattern class for such a pair and (optionally) the orientation as an additional flag of an annotation.

In conclusion, the Kunisch taxonomy is a very good starting point for our first prototype, as it constitutes an encompassing structure of geometric patterns with many examples. To the best of our knowledge, this is the most encompassing survey of geometric patterns of classic vessels. However, for long-term application and larger-scale annotation, we want to extend this existing pattern class structure or build up a new taxonomy altogether.

6.4.4 Applications

An encompassing annotated dataset not only allows for the training of dedicated pattern classifiers by providing labeled image pattern pairs but also enables a variety of other applications.

That is, the structured annotation generated with our system could be leveraged for various search and retrieval applications relevant to archaeological research. Examples include, but are not limited to: searching for objects exhibiting a certain pattern; searching for objects exhibiting a certain combination of patterns; or searching for objects exhibiting a certain pattern at a specific location on their surface.

With dedicated analysis methods, deviations in shape or location can be unveiled, leading to new discoveries and insights. Registering the individual occurrences of a certain pattern allows us to examine the intra-class variability, identify outliers, and even determine a prototypical pattern representation. Based on the location labels, correlations between the pattern classes and surface regions can be revealed.

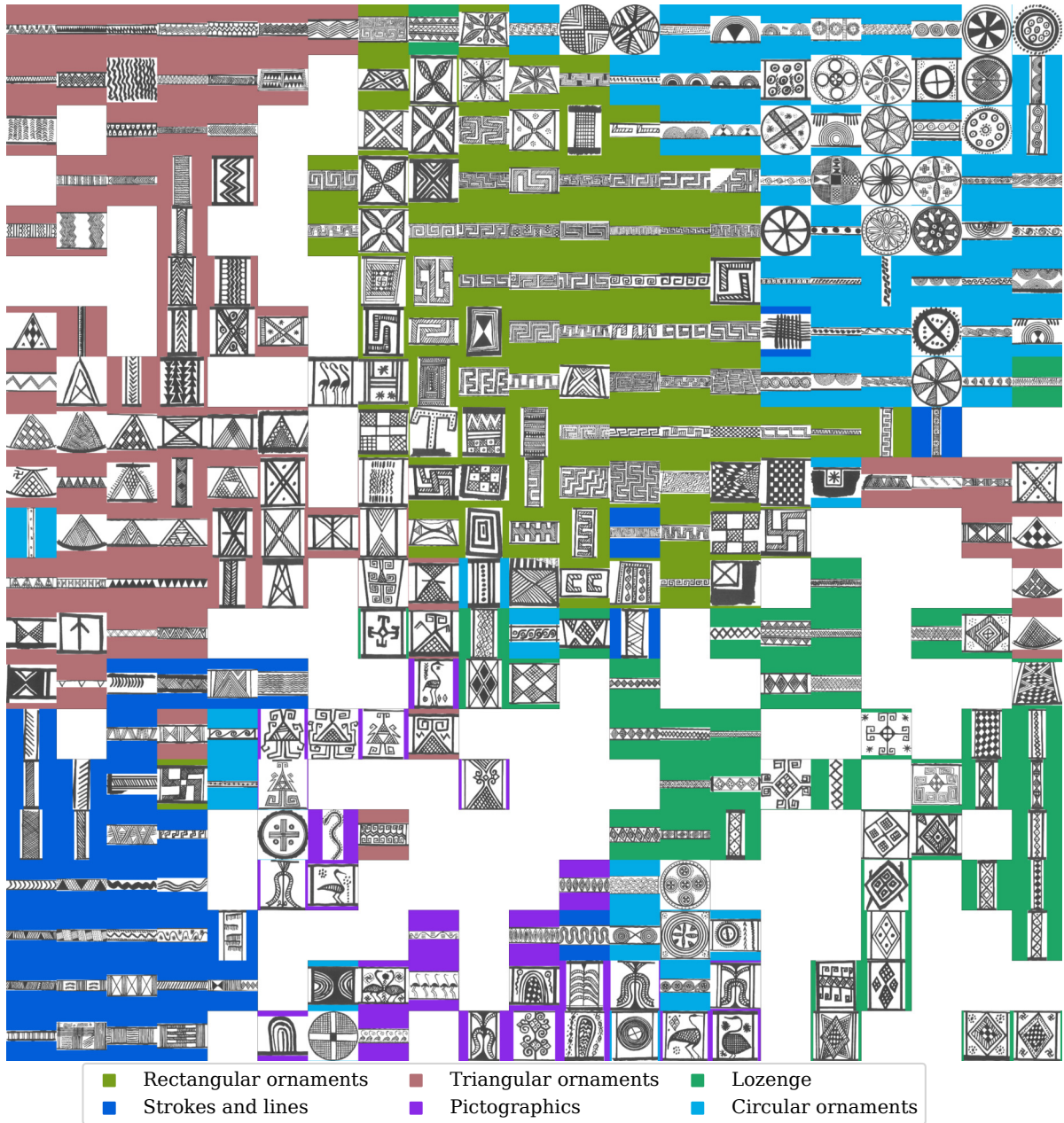


FIGURE 6.16. Visualization of the Kunisch patterns collection in a SOM based on FVs obtained from our CNN. The background colors indicate the affiliation to the major pattern groups.

A structured dataset also allows for visualization and (interactive) exploration. A huge number of methods have been studied to explore multimodal data at all different levels of visual granularity, supporting a broad overview down to close-up. *E.g.*, for the overview visualization of – potentially huge – repositories, visual cluster analysis methods like Self-Organizing Map (SOM) are an established visualization technique. This unsupervised machine learning technique generates a two-dimensional representation of high-dimensional input data. For a simple experiment,

we apply the approach by Barthel *et al.* [57] to the pattern drawings from Kunisch and obtained the map depicted in FIG. 6.16 based on the features obtained from the CNN (SEC. 6.4.1.3). Note that for illustrative purposes, we randomly selected just about half of the entire set of patterns, which were selected at random. The different background colors reflect the different major pattern groups. It can be observed that even though the CNN is unaware of this coarse classification, very similar clusters are formed where visually similar patterns generally appear close to each other. Other concepts [50, 56] also support the interactive exploration of data and allow the user to investigate selected subsets of the data in more depth and along specific traits in order to reveal correlations.

6.5 Semi-automated Creation of archaeological Drawings

The creation of graphical annotation of vase paintings – drawing in archaeological terms (FIG. 6.17 bottom) – is a crucial task in the research of ancient pottery. They improve the readability of a painting and enable its in-depth analysis and interpretation, such as the painter’s style and used manufacturing techniques. Manual creation with tools like tracing paper requires access to the physical artifact and is oftentimes neither feasible nor allowed due to the fragility of artifacts. Both 3D scanning and computer-aided unwrapping methods (SEC. 3.1) have considerably facilitated this task (FIG. 6.17 top) since they allow a domain user to annotate a painting directly on top of a digital planar surface representation. To date, the annotation is conducted mostly manually, for a domain expert is required to distinguish legit painting from various noise sources like fracture lines, signs of erosion, or worn-off and missing surface parts. Therefore, a mere edge image of the planar surface will not suffice (FIG. 6.17 middle). Still, manual annotation can take up to several hours, even for rather small vessels (FIG. 6.18) [74]. Consequently, repetitive ornamental patterns, which often take up large portions of vase paintings, are not annotated graphically but only described textually. Nonetheless, they are an integral part of the surface decoration and are important cues for attributing a pottery object to a painter or workshop and, overall, for dating the artifact.

In the following, we present a concept that greatly facilitates the graphical annotation of such ornamental bands using an interactive query-by-example approach, making it feasible for archaeologists to include them in surface drawings. To this end, we apply a similarity detection that leverages two constraints that commonly apply simultaneously to ornaments within an ornamental band: (i) an ornament is similar to a provided primitive and (ii) several of such occur at the same position along the object’s rotational axis.

6.5.1 Concept

Our approach takes as input a textured 3D model of a painted vase and outputs an image of an unwrapped projection of the object surface (henceforth referred to as *canvas*) exhibiting a



FIGURE 6.17. Top: The spherically unrolled 3D model of a Corinthian aryballos (*KFUG G26*, c.f. FIG. 5.1), exhibiting lavish vase paintings. The outlines of those paintings obtained naïvely with a Canny edge detection [259] (middle) are too erroneous and noisy for documentation purposes and cannot compare to a detailed manual annotation (bottom) where surface deficiencies are filtered out, and noisy parts reconstructed based on domain knowledge.

binary depiction of all extracted ornament groups. The group-wise detection and extraction are supported by a user operating on an interactive depiction of the projected vessel surface. We choose a projection that widely minimizes both distance and angular distortions [77]. The results closely resemble the surface as they would be drawn by an archaeologist, while it also allows for robust similarity detection in surface areas of high curvature.

In our proposed workflow (FIG. 6.19), the following four steps are executed repeatedly until all ornamental elements are extracted: ① the user selects one well-preserved ornament primitive from an ornamental band with a scalable selection rectangle; ② a self-similarity detection (SEC. 6.5.2), yielding a set of candidates (blue circles), is computed in the background; ③ false

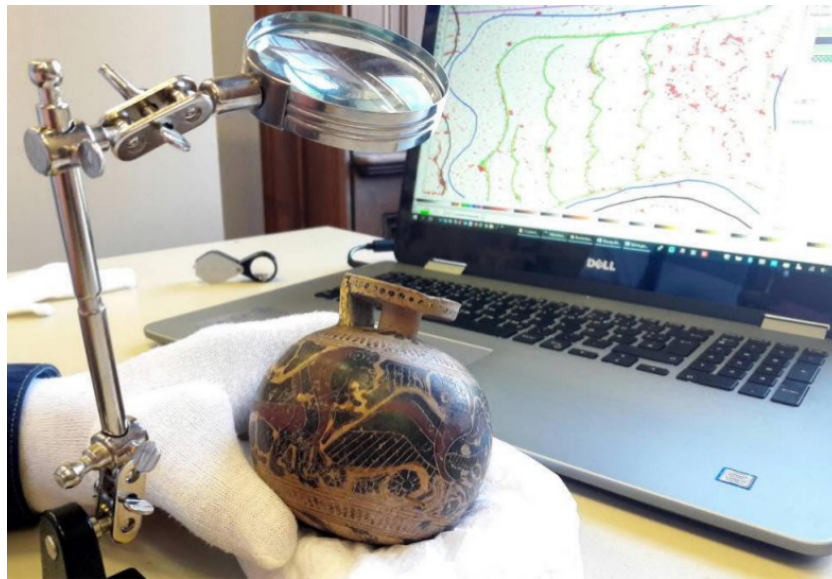


FIGURE 6.18. The manual creation of a vase painting's archaeological drawing – here given with the example of the Corinthian aryballos *KFUG G26* – is a labor-intensive process that involves the alternating investigation of intricate details on the original artifact and the tracing in a graphics editor. The result of this effort is shown in FIG. 6.17, bottom. (Image source: Stephan Karl)

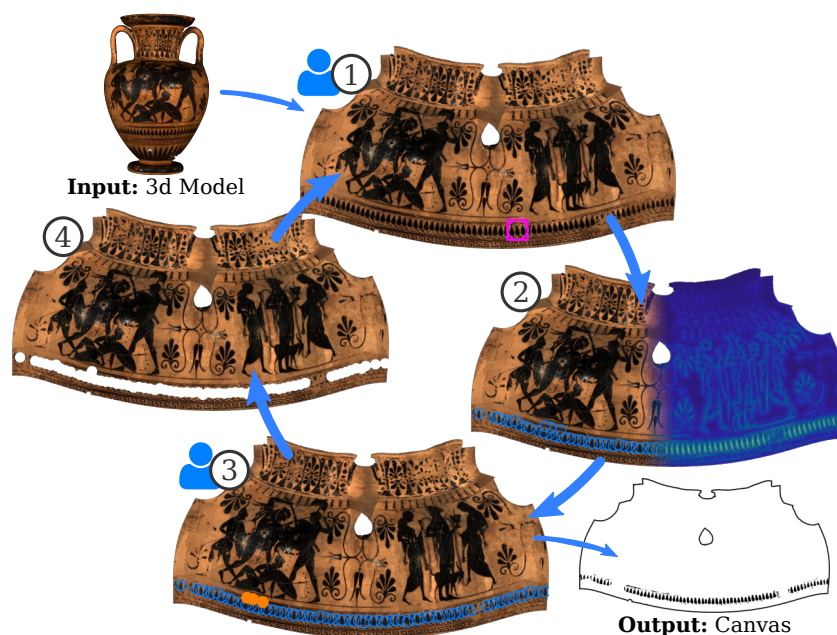


FIGURE 6.19. Workflow overview with a 3D model of the Attic black-figured amphora *KHM IV3600* as input, a canvas as output, and the iterative steps of ① query selection, ② self-similarity detection, ③ result refinement and ④ surface subtraction. A blue symbol indicates steps requiring user interaction.

positives (orange) are discarded by the user before a binarized representation of the refined selection is written to a separate canvas; and ④ the displayed vessel surface is updated by

removing surface parts containing the remaining selection. The canvas resulting from the extraction of all repetitive ornaments can be used as a basis for a complete annotation of the painting showing several figures.

6.5.2 Self-similarity Detection

With a sliding window approach, we compute the pixel-wise correlation of a query patch, given by the user's selection, and an equal-sized patch of the projected surface. The resulting confidence map can be employed for making proposals (candidate regions similar to the query). As we tailor our workflow specifically to the annotation of ancient Greek pottery, we can leverage the assumption of a bichrome colorization (*c.f.* SEC. 1.1). *I.e.*, we transform the colored surface into a binary representation prior to any detection efforts. In contrast to photographs, scanned 3D models often exhibit regional intensity differences due to locally varying lighting conditions of the scanned vessel surface. We use the locally adaptive thresholding method by Sauvola *et al.* [260] (FIG. 6.20 top left) to account for these impairments. Using a binarized input has the significant advantage that the correlation between query and image can be computed with logical conjunction instead of the more complex similarity computation between colors. That is, we can employ a linear filter that convolves the binary surface $B \subset \mathbb{B}^{w \times h}$, of width w and height h , with the $w_q \times w_q$ query (kernel) $\omega \subset \mathbb{B}^{w_q \times w_q}$, giving a normalized pixel-wise confidence

$$(6.5) \quad C_k(i, j) = \frac{1}{w_q^2} \sum_{\substack{i' \in [0..w_q) \\ j' \in [0..w_q)}} \phi(i', j') \omega(i', j') B(i + i', j + j'),$$

for the pixel at $\langle i, j \rangle$. $\phi = 1$ is an optional weighting term. For sufficiently large kernels (about $w_q > 10$), which is usually the case for this application, it pays off to leverage the *convolution theorem*, which can reduce the quadratic time complexity of a naïve convolution to $\mathcal{O}(n \log n)$. Aforesaid states that the convolution of two functions in real space equals the product of their respective Fourier transform in Fourier space [261]. More specifically, the image is divided into blocks of $w_q \times w_q$ pixels before the kernel is convolved with each block separately. Those convolutions yield results of dimensions $2w_q \times 2w_q$, which are added back together in an overlapping fashion, which is referred to as *overlap add*.

The detection has to be rotation-invariant, for relevant ornaments can occur in a rotated manner, either due to alternating orientations in an ornament band, or due to distortions caused by the preceding surface flattening step. This invariance is achieved by calculating the confidence for a set of K queries, with the k -th query exhibiting a rotational transformation equal to an angle of $2\pi k/|K|$, *w.r.t.* the original ($|K| = 90$ radial bins are used in all experiments). As the comparison of a rectangular selection is not reasonable in the rotated case, the query is masked, yielding the contents of the rectangle's inscribed circle exclusively (FIG. 6.20, pink). This is also

accounted for in the convolution (EQ. (6.5)) by setting ϕ to

$$(6.6) \quad \phi_{i'j'} = \begin{cases} 1 & \text{if } \|\langle \frac{w_q}{2}, \frac{w_q}{2} \rangle - \langle i', j' \rangle\| \leq \frac{w_q}{2} \\ 0 & \text{otherwise.} \end{cases}$$

Let $C \in \mathbb{R}^{m \times n \times K}$ be the stacked confidences (EQ. (6.5)) with $[C]_{ijk}$ as the confidence at pixel $\langle i, j \rangle$ and radial bin $k \in [0..K)$. The accumulated sliding window confidence $C_{sw} \in \mathbb{R}^{m \times n}$ (FIG. 6.20 top middle) is given by the pixel-wise maximum confidence $[C_{sw}]_{ij} = \max_{k \in [1..K)} [C]_{ijk}$.

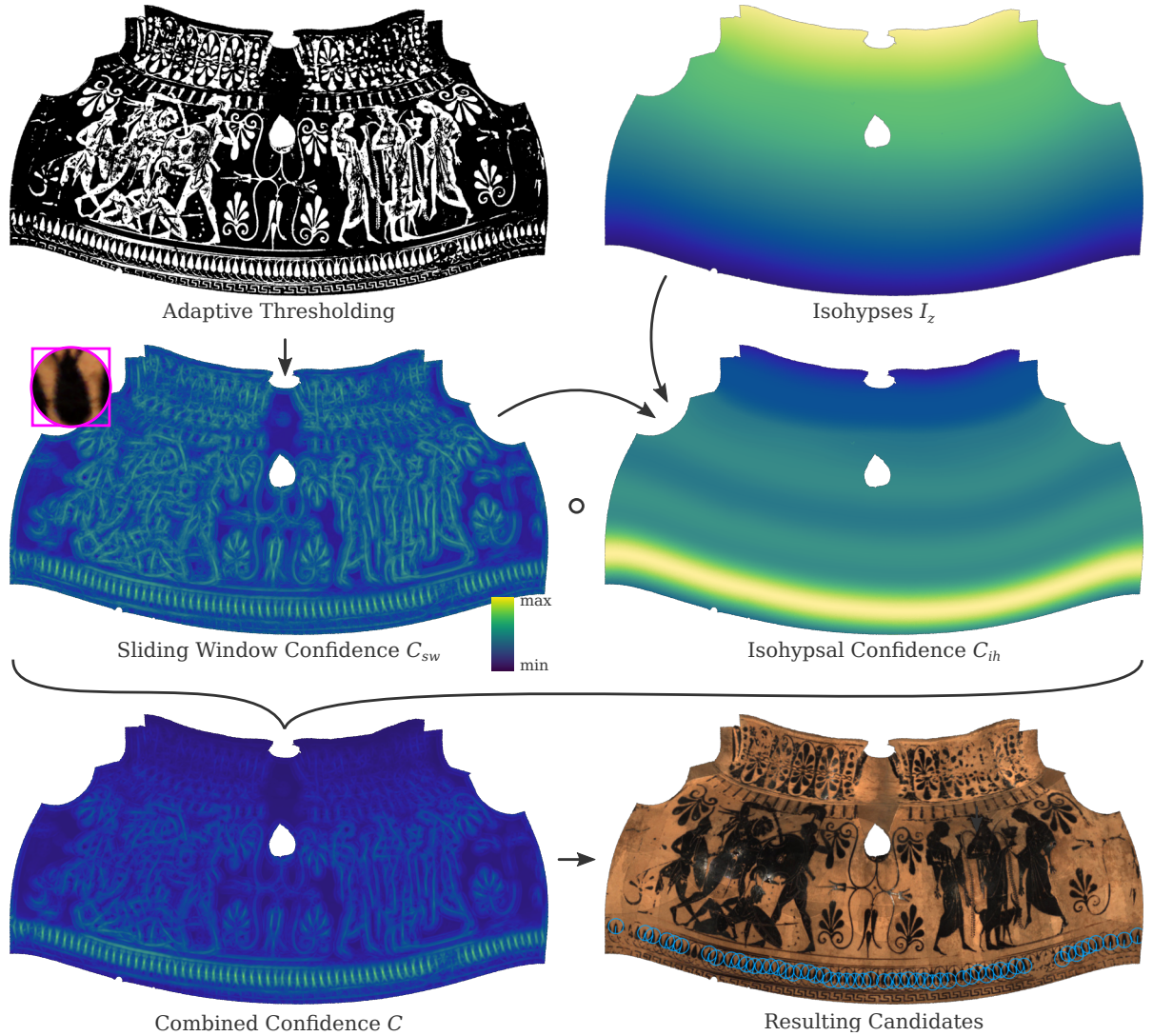


FIGURE 6.20. The unwrapped surface is subjected to an adaptive thresholding. Over the result we compute the probability C_{sw} of an area exhibiting a given reference pattern using a sliding window approach. Correlating this information with the areas of equal height (isohyps) I_z yields the isohypsals confidence C_{ih} . An element-wise multiplication of both confidence maps gives the resulting combined confidence C from which candidate regions are obtained using the NMS.

Entity Extraction from Confidence Map. The sliding window correlation results in a pixel-wise similarity of the respective locations to the query. Yet, the desired outcome of a detection pipeline is a finite set of candidate regions with very limited overlap. To this end we employ the NMS, outlined in ALGORITHM 1. In our case, the initial bounding boxes $\mathcal{B}$ are the boxes with dimensions $w_q \times w_q$ centered around image pixels whose confidence score $C(i, j)$ is above a threshold t_C . For the tested inputs, $t_C = 0.5$ proved to be a good choice. Together with the $\mathcal{B}$ and their respective confidence scores, obtained from C at the center point locations, we provide the IoU threshold of $t_{NMS} = 0.3$. An application of the NMS with these settings to the Amphora surface is given in FIG. 6.20, bottom right, with the final bounding boxes visualized as blue circles inscribed to them. The lines between the circles and their respective centers indicate a match's orientation *w.r.t.* the query.

Refinement with Isohypses Map. Queries with atomic shapes, *e.g.*, the blossom in FIG. 6.20, yield several high-confidence peaks strewn around across the whole confidence map. Due to its simplicity, similar patterns frequently appear in various parts of the surface, resulting in numerous false positives after the NMS. For the blossom query, 3 of the 68 detected boxes (4.4%) belong to this category. We refine the capability of our repetitive ornament detection by leveraging the assumption that repeating ornaments are arranged in ornamental bands around the rotational axis of the solid of revolution. With a mapping between the pixels of the flattened surface $S \in \mathbb{N}^2$ and the vertices of the 3D model $V \in \mathbb{R}^3$, we can derive the position of a surface point $P \in S$ along the rotational axis (Z-axis) of the solid of revolution. With $I_z : \mathbb{N}^2 \rightarrow \mathbb{R}$ as the map of z positions of the surface pixels, we are able to determine surface areas of equal height (isohypsal areas, FIG. 6.20 top right). An isohypsal confidence map C_{ih} , of the same dimensionality as C_{sw} , determines the probability that a surface region is part of an ornamental band exhibiting the query ornament. To this end, we multiply the isohypsal mask $H_z = \mathcal{N}(z, \sigma^2)$ of a normal distribution around a specific z ($\sigma^2 = \text{const.}$) with $\phi_z = \sum_{i=0}^{|hits(H_z)|-1} C_{sw}(hits(H_z)[i])$, a value relative to the number of hits and their respective confidences in the masked region. $hits(H_z)$ returns the center points of the detection boxes obtained with the NMS, which lie within H_z . The isohypsal confidence $C_z = \sum_{z=0}^{|M|-1} H_z \phi_z$ (FIG. 6.20 middle right) is defined as the accumulation of the masked confidences for individual z positions, with $M = \{0, 2\sigma^2, 4\sigma^2, \dots, 1\}$ as the set of mean values. The (element-wise) combined confidence $C = C_z \circ C_{sw}$ (FIG. 6.20, bottom left) is the refinement of the initial sliding window confidence C_{sw} and results in significantly fewer false positives when used as input for the NMS (FIG. 6.20, bottom right).

6.5.3 Experimental Results

With our prototype implementation, we conduct an evaluation of the annotation pipeline on the Amphora object *KHM IV3600*. To this end, we selected five queries of ornaments with varying size and complexity (FIG. 6.21 left), whose detected similar surface parts are drawn to a canvas

(FIG. 6.21 right). Step ③ the selection of false positives is omitted since a working UI for this task is not part of our preliminary prototype implementation. Yet, it can be observed that – even without step ③ – the combination of sliding window and isohypsal confidence values results in several correct detections for each query. It appears the areas with missing detections generally correlate with areas exhibiting strong abrasion artifacts resulting in partially (or altogether) missing ornament paintings. This is especially the case with the tear shape (red) and the meander shape (violet) at the lower half of the surface. It is also an issue with the leftmost occurrence of the blossom shape, which is almost entirely missing. However, it also has only four entities, and the similarity detection capabilities generally decline with both, increasing query complexity as well as a decreasing number of repetitions.



FIGURE 6.21. Flattened surface of *KHM IV3600* (left) with different ornament selections marked by colored circles and the resulting drawing on the canvas (right) with the same color coding.

Even though our approach is optimized for vessel surfaces exhibiting a bichromatic colorization, *e.g.* Greek pottery, it is also (to some degree) applicable for polychromatic surface paintings, observable on pottery of other ancient civilizations. FIG. 6.22 shows the outcome of the similarity detection and NMS for two pre-Columbian pots (SEC. 2.2.5). While our method works well for the first object, where all of the ornament's repetitions are correctly detected, it reaches its limits for the second object with a success rate of approximately 50% and also some false positives. This can be attributed to the suboptimal binarization of this surface as well as to the fact that the ornament in question also appears in a similar fashion as an atomic building block of the more sophisticated face ornament. Nonetheless, the presence of multiple ornamental bands containing the query ornament is visible in the isohypsal confidence map, allowing to discard of many of the false positives present in the sliding window confidence map.

6.5.4 Limitations and Outlook

The presented detection pipeline is comprised of a collection of (mostly) off-the-shelf image processing techniques. While they do the job for the targeted application of ancient Greek pottery, adoptions are necessary to achieve generalizability, scalability, efficiency, and robustness. Regarding the generalizability, we refer to the bichromatic colorization assumption, which does

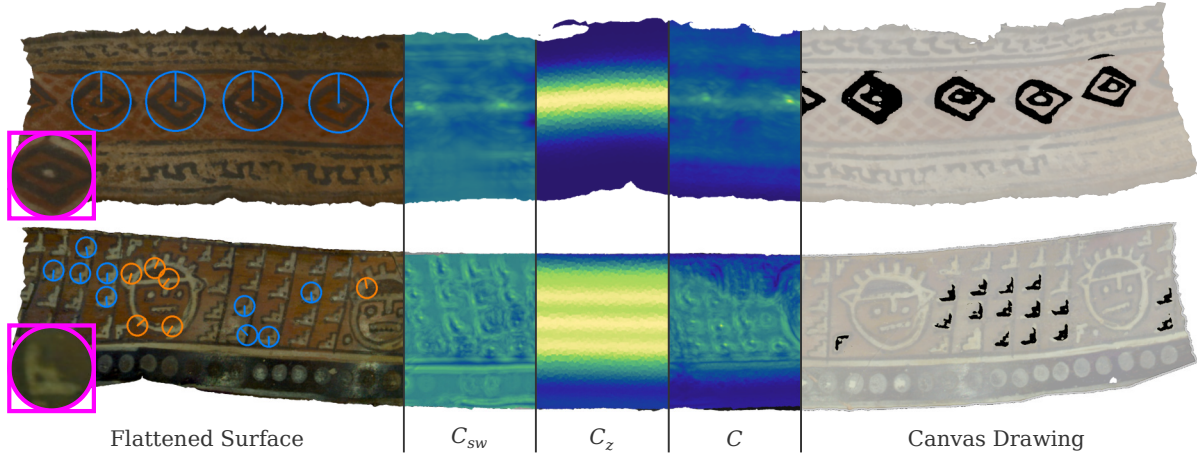


FIGURE 6.22. Filtered candidates of the similarity detection (blue circles for true positives and orange circles for false positives) for provided queries on two Peruvian pots and illustrations of the sliding-window confidence C_{sw} , isohypsal confidence C_z , their combination C , and the resulting extracted ornament drawings.

not hold for vase paintings in general, as can be seen, for example, in FIG. 6.22. To support polychromatic paintings the binarization step has to be changed to a semantic- or even an instance segmentation (*c.f.* SEC. 3.2). If ornaments within a band appear with varying sizes and in a mirrored fashion, this needs to be also factored into the sliding window detection. While it is possible from a technical standpoint – *e.g.*, in a similar fashion to the rotation invariance with a series of rotated kernels, differently scaled and mirrored versions of the convolution kernel could be used – the computation complexity increases exponentially (even though the convolution runs efficiently) with every new dimension of invariance. This merits the application of sampling-based detection approaches or techniques relying on deep learning [222]. Such changes would also boost the computation times in general, which are acceptable for first experiments, with a couple of minutes on commodity hardware for a high-resolution input, but a real-time interactive annotation or the annotation of a large dataset this is not acceptable.

The proposed detection with the isohypsal refinement is not only applicable for the generation of archaeological drawings as in our pipeline but could also be used to extract a database of ornaments from a repository of textured 3D vessels. Such could be used for a content-based retrieval allowing us to search for similar ornamental depictions on other objects or to group ornament entities by shape, as shown by Karl *et al.* [74] for *KFUG G26*. Further down the line, this also allows deriving an ornament band’s inherent grammar, which can be used to detect and reconstruct missing or eroded parts (CHAPTER 8). Yet, to this end, the self-similarity detection needs to run purely unsupervised, and thus it has to be – even though occasional false positives are no concern in a retrieval setup – more robust in general.

6.6 Conclusion

Within this chapter, we present three different contributions pertaining to the subject of retrieval by ornamentation. First, a benchmark dataset that allows the objective comparison of different recognition techniques for patterns. *W.r.t.* to the tasks introduced in SEC. 6.1, this particularly addresses **T0** and **T1**. Second, we present a system for the online annotation and classification of Geometric pottery, which falls into the same category of tasks. Finally, we present an approach supporting the annotation of pottery surfaces with a self-similarity detection procedure in the background. This system serves the documentation **T3** of the ornamental parts of vase paintings. Even though **T2**, the (statistical) analysis of ornaments was not explicitly covered within the presented works, it has been the focus of a recent Master's thesis [230] based on the presented benchmark dataset. Using artifacts from different cultures and epochs for evaluation purposes underlines the generalizability of the proposed methods. While decent results are achieved with the semi-automatic annotation and the classification of Geometric pottery, the problem seems to be far too complex for fully automated processes. Hence, we believe that, even with further advancements in ML techniques and the increasing availability of labeled data, supervision by a user will remain a crucial element to any system.

Part III

Visual Analytics

INTERACTIVE EXPLORATION AND VISUALIZATION

Contents

7.1	Introduction	142
7.2	Related Work	142
7.3	Linked Views Visual Exploration System	144
7.3.1	Domain Analysis Tasks	145
7.3.2	Concept	145
7.3.3	Analytical and Interactive Support	150
7.3.4	Results and Evaluation	154
7.3.5	Discussion Limitations and Open Research Questions	156
7.4	Conclusion	157

In the previous chapters, we employ different notions of object similarity for sorting a collection of records *w.r.t.* a certain query item. The same metrics can, however, be used to introduce an ordering among the collection’s records. Such is insofar relevant as domain researchers often do not want to search for objects similar to a given reference but aim to discover (previously unknown) similarity clusters within a collection. Or with the words of Martínez-Carrillo [59]:

Data archaeology is a skilled human task, in which the knowledge sought depends on the goals of the analyst, cannot be specified in advance, and emerges only through an interactive process of data segmentation and analysis.

Yet, in contrast to searching where a one-dimensional result vector is the preferred way of displaying a retrieval result, this is a poor choice for illustrating overall object similarity in a dataset. The research area of visualization is concerned with finding an optimal solution for this purpose.

This chapter is based on:

- Lengauer, S., Komar, A., Karl, S., Trinkl, E., Preiner, R., Schreck, T., “Visual Exploration of Cultural Heritage Collections with Linked Spatiotemporal, Shape and Metadata Views”, in *Vision, Modeling, and Visualization*, The Eurographics Association, 2020. DOI: [10.2312/vmv.20201196](https://doi.org/10.2312/vmv.20201196)

7.1 Introduction

The analysis of CH artifacts in terms of their historical context is of fundamental importance for archaeological research. To this end, finding and understanding relations between artifacts regarding traits of different inherent and derived modalities, *e.g.*, spatial (findspot, provenance), temporal (dating), or categorical (shape, material, style), is a crucial task that often stands at the beginning of gaining new insights. Thanks to the effort of Digital Humanities, more and more artifacts are digitized and collected in databases together with the relevant metadata that encodes these traits and allows for efficient queries for relevant comparable artifacts using classic ‘search by metadata’ techniques. While this supports the search for artifacts with a given specific set of traits, domain experts often gain important insights by analyzing large volumes of artifacts as a whole, *e.g.*, by investigating the relations between clusters based on different trait modalities. However, classic search-based interfaces to digital repositories seldomly allow archaeologists to analyze a suitable overview of the whole volume of a collection’s objects, limiting the exploration of potentially interesting relations among artifacts across different modalities. Moreover, in practice digital repositories typically contain artifacts with missing, unknown, or uncertain metadata, introducing an additional challenge to the assessment of connections between groups of artifacts based on linked modalities. Finally, much of the information necessary for revealing such connections is contained within artifact-related traits like their shape, painting style, or similar. *I.e.*, it is available only latently in the form of object depictions, instead of being encoded in a low-dimensional metadata domain to search in.

We propose a design for an exploration system, utilizing the potential of digital collections for a task-oriented analysis and exploration of CH data. The system addresses the above-mentioned challenges by supporting the domain user in the analysis of large digital artifact collections and the assessment of significant relations between artifacts based on multi-modal traits. Related designs from the CH domain and others are discussed in SEC. 7.2 before our concept is outlined in detail in SEC. 7.3. The chapter closes with a summarization of our contribution and a few general thoughts on visualization for the CH domain (SEC. 7.4).

7.2 Related Work

To date, many works support domain research via interactive visualization. There is also a growing number of systems, specifically for analyzing CH data [107]. The majority relies primarily

on visualization of text and metadata including spatiotemporal, relational, image, and shape information. However, fewer consider these types of data in a tightly coupled way.

Visualization Techniques by Modality. Appropriate visualization concepts have been studied based on the modality of the data at hand. The most widespread type of data is textual information. Regarding this type of data Moretti [262] coined the term *distant reading* which refers to the visualization of texts with, e.g., graphs, maps, or trees, as opposed to the traditional form, referred to as *close reading*. Jänicke *et al.* [263] present a survey on respective techniques.

To visualize sets of images, the optimal overall layout for arranging the images is usually the crucial question. Brivio *et al.* [264] propose a Voronoi-based layout to visualize photographic campaigns in CH. Glinka *et al.* [265] show the potential of details-on-demand techniques for the exploration of large CH collections including images, keywords, and textual data. They use a zoomable timeline visualization to combine a “distant-viewing” and “close-viewing” mode in an integrated exploration system. Windhager *et al.* [107] give an extensive overview of both the different data modalities relevant to CH as well as appropriate visualization techniques for each of them.

Multimodal Visualization Systems. Graph-based visualization techniques are frequently used in cases where corresponding elements have inherent relationships with one another. A graph is given by a set of nodes and edges and is commonly used for the analysis of structures and grouping of nodes and/or edges (Compound Graph Visualization) [266]. Further application areas and examples of graph-based visualizations are given in dedicated surveys [267, 268]. Regarding applications for CH, Novak *et al.* [43] apply graph layouts to show historical social network relations together with their temporal aspect in a separate timeline. Preiner *et al.* [44] combine these aspects into an integrated network visualization on top of a height field displaying the temporal aspect. Network visualizations have also been used for visualizing abstract CH artifact characteristics, like vessel shape similarity by Van der Maaten *et al.* [45], or for displaying similarities across Mayan inscriptions [46, 47]. Tortora *et al.* [48] use graph views to support archaeologists in finding new correlations from ontology-driven metadata and Coburn [49] proposes a map view of artifacts that can reveal semantic relationships between artifacts via visual links.

Other approaches illustrate multiple properties simultaneously in a high-dimensional space, such as the approach by Windhager *et al.* [50], who show the space-time attributes of CH artifacts within a three-dimensional PolyCube system. By connecting different visualization techniques, like map, set, and network visualizations, they reveal spatial, categorical, and relational collection aspects. Simon *et al.* [51] introduced Peripleo, an open-source tool for exploring the geographic, temporal, and thematic composition of distributed digital collections. To the same end, *multi-view* approaches, have seen widespread use as noticeable from the survey by Roberts *et al.* [269]. Relevant previous works in multi-view and linked visualizations include the Vis-Trails framework

[270], which allows connecting individual views together in an interactive environment. Furthermore, the Improvise framework [271] is also influential in defining linked view arrangements for data exploration. Steed *et al.* [272] present the CrossVis system, a combination of several interactive views of multivariate data. That is, the main view, showing parallel coordinates, is supported by scatterplots, correlation visualizations, and image views.

Positioning of Our Work. The visualization concept proposed in the following does not introduce novel visualization techniques, but it relies on established concepts (multi-view, graphs, maps, timelines, parallel links) which are combined in an ingenious manner to allow for the concurrent exploration of data collections along different object traits. While only three traits are visualized in a purposely created prototype implementation, the advantage of our system is that a potentially arbitrary number of traits can be supported, as opposed to many other concepts which are limited in this regard due to their design.

7.3 Linked Views Visual Exploration System

We employ multiple linked views revealing the localization, distribution, and concentration of artifacts along different attribute domains, as well as their connections across them. For a proof-of-concept, we focus on the three attributes *(i)* provenance, the location an artifact originates from; *(ii)* dating, the epoch/time an artifact supposedly stems from; and *(iii)* shape. These properties are of particular interest to pottery research [59, 141]. To provide an overview of artifact records based on their shape, we propose an interactive graph-based representation that structures artifacts based on a custom shape-similarity metric. We support adaptive levels of visual granularity by introducing multi-object previews tailored to shape aggregations. To this end, we adapt an established graph sparsification algorithm for locally varying link densities to work for fully-connected undirected weighted graphs. Moreover, we introduce a new cross-correlation measure for graph layouts, allowing the automatic detection and highlighting of prominent artifact correlations across different modality domains. Said reveals potentially interesting connections between relevant artifact subgroups. Views on different modalities naturally incorporate missing and uncertain data of individual artifacts through respective visual cues.

The proposed design concepts are tailored to an interactive visual exploration system for artifact collections that exhibit shape and spatiotemporal metadata. Nonetheless, the system's design can be readily extended to support additional object traits. Albeit not limited to this domain, our prototype implementation is evaluated on a collection of CH objects. Additional challenges, specific to this domain, are the uncertainty of attributes, *e.g.*, dating, as well as missing attributes for some items. For the design and the evaluation, we rely on a set of meaningful tasks defined by domain experts. The potential of the approach is demonstrated using a Greek pottery collection, revealing relevant correlations regarding the historical context of ancient pottery in an intuitive and comprehensible manner.

In the following, we discuss the mentioned domain analysis tasks (SEC. 7.3.1) before the proposed design is shown in-depth in SEC. 7.3.2. SEC. 7.3.3 explains how various analysis and interaction tasks are enabled through customized components, and SEC. 7.3.4 discusses the insights gained from a collaborative walkthrough together with domain researchers. Finally, SEC. 7.3.5 concludes the section by stating limitations as well as open research questions.

7.3.1 Domain Analysis Tasks

In discourse with domain researchers, we gained insight into their research workflows, based on which we determine a set of relevant tasks for an archaeological exploration tool. The tasks are categorized as being of single-modal (**T1-T4**), and cross-modal (**T5-T6**) type and include:

- **T1** Gaining an overview of the geographic distribution of the artifacts' provenances;
- **T2** Examining patterns in the artifacts' temporal distribution;
- **T3** Gaining an overview of the diversity and variability of artifact shapes;
- **T4** Examining artifacts at a dynamic level of visual granularity, down to close-up;
- **T5** Revealing connections between spatial, temporal, and shape-based distributions; and
- **T6** Evaluating salient connections determined by a cross-correlation measure.

These tasks serve as a guideline in the design and the evaluation of the proposed visualizations.

7.3.2 Concept

Coherently displaying two or more dataset traits in an integrated visualization is a non-trivial problem, which is even aggravated if these traits comprise different types of attributes (categorical, geographic, and abstract in our case). *I.e.*, we decide to diverge from the idea of visualizing all traits in a single component towards a concept that uses different views for the various modalities. This is a known concept for visualizing multiple traits of attributed data and is referred to as *multiple view* approach by Roberts *et al.* [269]. The advantage is that tailored visualization techniques can be employed for the individual traits as well as that it permits the user to systematically explore inter-modality characteristics. On the downside, multiple views have the disadvantage that the user's attention is focused on one view at a time, making it hard to grasp the overall structure of the data and spot intra-modality correlations. Common ways to overcome this issue are coordinating, synchronizing, or linking actions between individual views. For the problem at hand, the latter, also referred to as *multiple linked views* [269], proved to be the most convenient. Aggregation and clustering methods facilitate the handling of vast object collections. We design the system such that individual artifacts are perceptible at the lowest level while the views become more and more abstract at higher levels. This is referred to as *visual*

granularity [107], and finding the appropriate level is an essential design objective for any CH InfoVis system. Our proposed system, dubbed Linked Views Visual Exploration System (LVVES), shown in FIG. 7.1, supports a dynamic level of visual granularity. *I.e.*, at all time, approximately the same number of previews is shown, but the level of detail changes based on user navigation interaction. In the remainder of the section, the individual views for our selected modalities are described in detail.

I.e., the Geographic Map Viewer (GMV) for spatial information, the Timeline Viewer (TV) for temporal information, and the Shape Similarity Viewer (SSV) for shape information. All of them have an underlying graph data structure $G(V, E)$ in common, where the vertices V (common across views) correspond to the displayed artifacts A and weighted edges E model the similarity between artifacts. These graphs are the basis for a *force-directed layout* [273], which dictates the positioning of an artifact (visualized with a preview) within a view through a set of forces acting simultaneously upon the vertices.

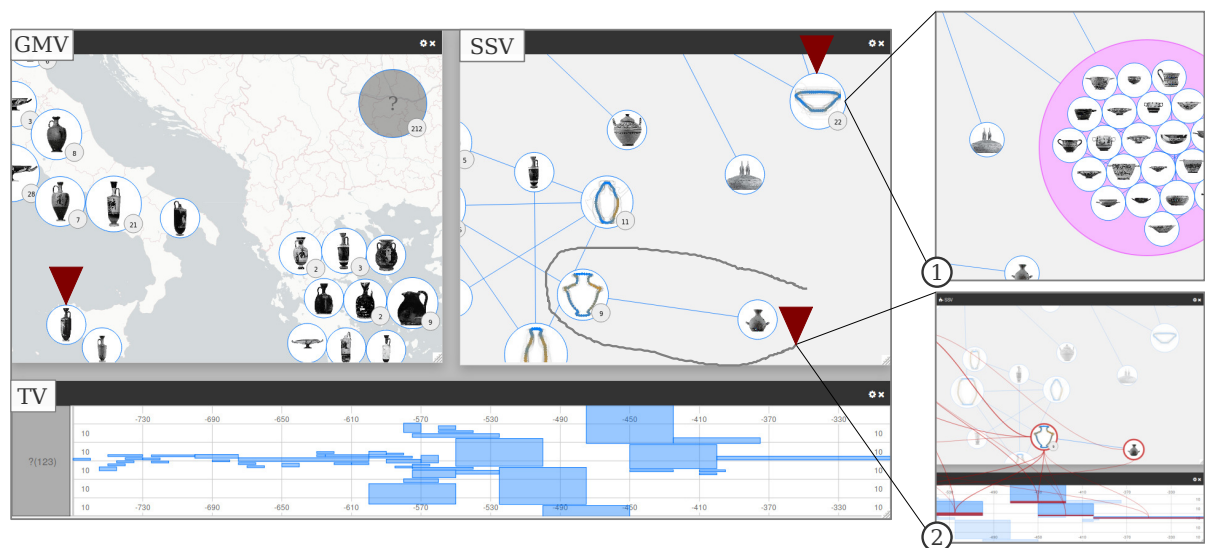


FIGURE 7.1. The Linked Views Visual Exploration System (LVVES) allows us to explore space (GMV), time (TV), and shape (SSV) modality in an integrated system with linked views. In each view, the artifacts are aggregated and clustered *w.r.t.* the respective modality using a dynamic degree of visual granularity, from overview to close-up ①. We introduce MOPs tailored to the SSV (*c.f.* FIG. 7.2), providing visual cues regarding the artifacts substituted by an aggregation. Intra-view cluster correlations are revealed using alluvial-flow diagrams ② spanning between views.

7.3.2.1 Geographic Map Viewer

Geographic maps are an established approach for visualizing spatial traits of CH artifacts [274]. As required in task T1, an artifact's origin within the GMV is indicated by a *preview*. If the preview stands for a single artifact, it is referred to as Single Object Preview (SOP) [107] and can display 3D models, images, text, or metadata. As all artifacts from our database have

associated photographs, we decided to display one representative photograph embedded in a circular marker as a preview (FIG. 7.1, GMV). If a preview surrogates multiple artifacts, we speak of a Multi-Object Preview (MOP) [107]. Arrangements like lists, grids, or mosaics are common representatives of this type. In the GMV, we display one randomly selected image from the respective set of artifacts as MOP in case of aggregation. Assigning an artifact to a specific point on a map requires a latitude/longitude coordinate pair. The provenance attributes associated with our artifacts, however, are given as place names with highly varying degree of geographic expansion, as findplaces are oftentimes only vaguely documented. *I.e.*, in some cases, these names refer to ancient cities, *e.g.*, Nola, while they can also refer to entire modern countries such as Italy or Greece. Obtaining a latitude/longitude coordinate pair for such a place name involves a process called *forward geocoding* and is offered by various web services. Nonetheless, a place name cannot be pinpointed to one exact location on the map but only to an area. We use the bounding boxes of these areas as positional constraints for the previews, meaning that a preview can roam freely within its associated bounding box. Note that these bounding boxes can intersect each other or even be contained within one another due to the large areas associated with some place names.

To get exact locations, as required for drawing the previews on the map, we use a force-directed layout with a gravitation force $f_g(v)$ pulling a vertex $v \in V$ towards the center of its respective bounding box with a constant strength. As soon as a vertex is located within its bounding box, this force becomes $\mathbf{0}$. With this attraction force alone, many previews of artifacts stemming from the same or intersecting areas are fully or partially overlapping each other. We address this issue twofold. First, a collision force f_{col} , with $f_{col} \gg f_g$, ensures that previews are separated sufficiently to avoid overlapping. Second, we aggregated vertices by common place name into surrogates, resulting in a much smaller number of vertices $V_{agg} < V$. The number of artifacts represented by a MOP – the cardinality $|A_{MOP}|$ – is visualized with a text label (FIG. 7.2) as well as a slightly larger preview radius r amounting to

$$(7.1) \quad r(|A_{MOP}|) = k \log(|A_{MOP}|) + r_{SOP},$$

with $k = 0.3$ as a scaling factor and r_{SOP} as the uniform SOP radius.

The absence of spatial information for some artifacts prevents their placement on the map. However, we want to indicate their presence regardless. On the one hand, every view should contain the same set of objects to allow for visual linking. On the other hand, it has been shown that the visualization of missing data can be crucial for decision-making [275]. That is, we consider all such artifacts by means of a ‘placeholder’ preview within whom all of them are aggregated. It is marked by a preview displaying a question mark and is always placed in the top right corner of the GMV (FIG. 7.1, GMV).

7.3.2.2 Timeline Viewer

Linked timelines, animation, superimposition, and space-time cubes are often used methods for displaying temporal information [274]. For the LVVES, we use a timeline visualization to this end, as it is the only one of the aforementioned, showing exclusively temporal information. We also want to incorporate cardinality information – how many objects stem from a particular epoch – into our viewer. We implement this by means of a two-dimensional viewer with the temporal information encoded along X -dimension and the cardinality information encoded along the Y -dimension. Yet, CH artifacts cannot be attributed to an exact point in time but only to a (possibly half-open) date interval due to uncertainties in the dating. Aggregating the artifacts by common time intervals leaves us with a set of blocks we can place in the TV. Although the expansion in X and Y dimension of such a block is unambiguous, the problem that many blocks are overlapping each other, due to overlappings in the corresponding date intervals, remains. This can be corrected by rearranging the blocks along the Y dimension such that they do not intersect each other. Finding a global arrangement that makes the optimal use of the available space is a variation of the *bin packing problem* and is combinatorial NP-hard [276]. However, we found an approximate solution to be sufficient for practical use. That is, we apply a force layout with two counteracting forces: (i) a vertical force f_y governing that the vertices align around a vertical center line; and (ii) a larger collision force f_{col} , with $f_{col} \gg f_y$, pushing blocks apart which are intersecting each other. FIG. 7.1, TV shows one possible result of this layouting concept. Note that the TV also features a labeled grid facilitating the orientation. Same as with the GMV, we face the problem that many artifacts are missing a temporal attribution. We account for those with a designated block in a secluded area to the left of the timeline.

7.3.2.3 Shape Similarity Viewer

The SSV provides an abstract visualization of a collection’s shape variety and variability. As proposed by Van der Maaten *et al.* [45], we assume that an artifact’s shape characteristics can be derived from its silhouette. To this end, we make use of an artifact’s associated images I_a , which are subjected to manual filtering to discard images showing only close-ups of surface details or related hand drawings. From the remaining images object masks are extracted (*c.f.* SEC. 3.2.1), from which the respective largest contour is obtained with the topological structure analysis algorithm by Suzuki *et al.* [207]. We quantify the similarity between two contours using the SCD (*c.f.* SEC. 4.4.3.1).

Let $F(a) = \{\mathbf{x}_{SCD}(c_{max}(I_a[i])) \mid i \in [0..|I_a|]\}$ denote the set of FVs for artifact a , with c_{max} returning the largest contour of the image’s object mask and $\mathbf{x}_{SCD}(c_{max})$ as the corresponding FV. We define the similarity d between two artifacts a_i and a_j as the minimum Euclidean distance $d(a_i, a_j) = \min_{(\mathbf{x}_k, \mathbf{x}_l) \in \mathcal{X}} \|\mathbf{x}_k - \mathbf{x}_l\|$ between all feature vector pair combinations $\mathcal{X} = \{\langle F(a_i)[k], F(a_j)[l] \rangle \mid k \in [0..|F(a_i)|], l \in [0..|F(a_j)|]\}$.

The similarity relations between all artifacts are given by the symmetric square adjacency

matrix $\mathbf{B} \in \mathbb{R}^{|A| \times |A|}$ with $b_{ij} = b_{ji} = d(a_i, a_j)$, which can be treated as a finite weighted graph $G^{SSV}(V, E)$ with vertices $V \hat{=} A$ and edges $E = \{\langle i, j, w_{ij} \rangle \mid 0 \leq i < j < |V|, w_{ij} \propto b_{ij}\}$. In the SSV, the grouping of artifacts based on their mutual shape similarity is done by a force layout where a link force $\mathbf{f}_l$ between all possible pairs of artifacts, e.g., $\langle a_i, a_j \rangle$ is proportional to the inverse of the corresponding edge weight $\mathbf{f}_l(i, j) \propto w_{ij}^{-1}$, assuring that similar artifacts are positioned close together while dissimilar ones are forced to stay further apart. Three additional forces act simultaneously upon all vertices: (i) a collision force $\mathbf{f}_{col}$; (ii) a strong repulsion force $\mathbf{f}_{rep}$, with a strength exponentially decreasing with distance, meaning that vertices not connected by a strong link force are pushed apart, resulting in a generally ‘cleaner’ overall layout; and (iii) a counteracting weak gravitation force $\mathbf{f}_g$ ($\mathbf{f}_g \ll \mathbf{f}_{rep} < \mathbf{f}_l \ll \mathbf{f}_{col}$) pulling all vertices towards a common origin and preventing weekly linked groups from drifting off. In addition to the previews at the corresponding vertex positions, links are drawn as straight lines (FIG. 7.1, SSV).

Graph Aggregation. With one vertex per artifact, the resulting graph layout will be huge, hindering the perception of the collection at a macro level. We want to provide a global overview by means of abstraction as in the GMV, and the TV, where said is achieved with an aggregation by common provenance and common date respectively. In the case of the SSV, we do not have a comparable finite set of common values as a basis for aggregation, but we combine vertices iteratively until a desired level of aggregation is reached. A greedy pairwise aggregation removes the edge $\langle \hat{i}, \hat{j}, w_{\hat{i}\hat{j}} \rangle = \min_{\langle i, j, w_{ij} \rangle \in E_{agg}} w_{ij}$ with the lowest weight $w_{\hat{i}\hat{j}}$ together with the vertices $\hat{i}, \hat{j}$ it connects. For the vertices a substitute is added as a surrogate that is connected with newly added edges to the rest of the graph. This process is conducted iteratively, with $V_{agg}^{(0)} = V$, $E_{agg}^{(0)} = E$ at the initial step and $V_{agg}^{(k)} = V_{agg}^{(k-1)} \setminus \{\hat{i}, \hat{j}\} \cup \{\cdot\}$, $E_{agg}^{(k)} = E_{agg}^{(k-1)} \setminus \{\langle \hat{i}, \hat{j}, w_{\hat{i}\hat{j}} \rangle\} \cup \{\dots\}$ at the k -th step, until $|V_{agg}|$ reaches a given cut-off threshold t_{agg} . The detailed algorithm outlining this process is featured in APPENDIX B. For t_{agg} we use a value equal to the mean number of previews in the GMV and the TV, to ascertain similar visual granularity across all views. As an additional visual cue the number of objects surrogated by an aggregated vertex (cardinality) is encoded into the preview size r_{MOP} (Eq. (7.1)).

Graph Sparsification. Initiating the force layout with $\mathbf{f}_l > \mathbf{0}$ between all vertex pairs results in a hairball-like structure due to the layout being heavily overconstrained. The process of thinning out these link forces while preserving link-induced accumulations at the same time is generally referred to as *graph sparsification*. The similarity graph G_{agg}^{SSV} has edges between all vertices (fully connected) with a total of $|E_{agg}| = |V_{agg}|(|V_{agg}| - 1)/2$ edges. This class of graphs is referred to as a *complete graph*. The goal of sparsification is to find a subset of edges $E_{sparse} \subset E$ such that $|E_{sparse}| \ll |E|$. The trivial approach is to remove a fixed number of edges with the highest weights [277], or to omit all edges exhibiting a weight above a globally defined threshold t_{sparse} , such that $E_{sparse} = \{\langle i, j, w_{ij} \rangle \in E \mid w_{ij} \leq t_{sparse}\}$. A critical flaw of this approach is that the

locally varying edge weights are not taken into account, resulting in the dissolution of some clusters while too many links are preserved in others.

Hence, we base our sparsification on an approach that aims at discarding inter-cluster edges while retaining intra-cluster edges. We utilize the method described by Satuluri *et al.* [277] (see Algorithm 2 in their work), which we adapt for complete weighted graphs. *I.e.*, instead of using the Jaccard similarity coefficient for the similarity between two vertices $i, j \in V$ we define a distance reflecting their agreement regarding the distance to all other vertices in V . Formally, it is given by

$$(7.2) \quad \text{Sim}_V(i, j) = \text{Sim}_E(E_i, E_j),$$

with E_i, E_j as the edges incident to i and j respectively, sorted by their weights, and

$$(7.3) \quad \text{Sim}_E(E_i, E_j) = \left[\frac{1}{\min\{|E_i|, |E_j|\}^2} \sum_{k \in [0.. \min\{|E_i|, |E_j|\})} |k - \text{indexOf}(E_i[k], E_j)| \right]^{-1}.$$

$\text{indexOf}(e, E)$ returns the index of the edge e in E . The algorithm used for obtaining G_{sparse} using this measure is outlined in detail in APPENDIX B. The advantage of this local sparsification method over the global one is that it has the tendency to preserve clusters within the graph structure since intra-cluster edges are much more likely to be preserved than inter-cluster edges.

Shape Preview Surrogate. Same as with the GMV, the SSV requires MOPs for aggregated vertices. Since this view deals exclusively with artifact silhouettes we conclude that such previews should also embed this information. To this end, we implement a glyph that displays the Euclidean mean of a set of silhouettes together with a color coding implying the local variances of silhouette shapes as depicted in FIG. 7.2. For a set of aggregated artifacts A_{agg} , associated with a MOP, this involves normalization and registration of all the representative silhouettes $H(A_{\text{agg}}) = \{c_{\text{max}}(I_a[0]) \mid a \in A_{\text{agg}}\}$. Note that only one silhouette per artifact is taken.

Let $\mathbf{s}_h$ be the vector of n_{SPS} points, sampled at equidistant positions along the registered silhouette $h \in H$. A Shape Preview Surrogate (SPS) is described by a point vector $\mathbf{s}_{\text{SPS}}$, initialized with the point-wise Euclidean mean of two silhouettes $\mathbf{s}_1$ and $\mathbf{s}_2$ with $s_{\text{SPS}}^{(0)}[i] = \frac{1}{2}s_1[i] + \frac{1}{2}s_2[i]$, $i \in [0..n_{\text{SPS}})$. All additional silhouettes are added online, meaning that the k -th silhouette vector $\mathbf{s}_k$ can be added with $s_{\text{SPS}}^{(k+1)}[i] = s_{\text{SPS}}^{(k)}[i] + k(s_k[i] - s_{\text{SPS}}[i]^{(k)})$, $i \in [0..n_{\text{SPS}})$. Additionally, we calculate the standard deviation of each point in $\mathbf{s}_{\text{SPS}}$. The glyph we create from this information consists of a series of dots representing the mean silhouette on a neutral background with the standard deviation encoded in color. Same as in the GMV, the size of a glyph in the viewer scales with its cardinality according to Eq. (7.1).

7.3.3 Analytical and Interactive Support

The visual connection across views is established via user interaction. The user selects a subset of the data in one of the views by a selection action, which involves, in many cases brushing or, in

our case, a selection lasso. The selected subset is then highlighted in all views by changing color or size. We take it one step further and also draw links between views in order to visualize the correlation of clusters and aggregations across modalities (SEC. 7.3.3.2). Besides this highlighting, we also provide the user with an automatically calculated ‘most significant’ selection based on a cross-correlation measure (SEC. 7.3.3.3).

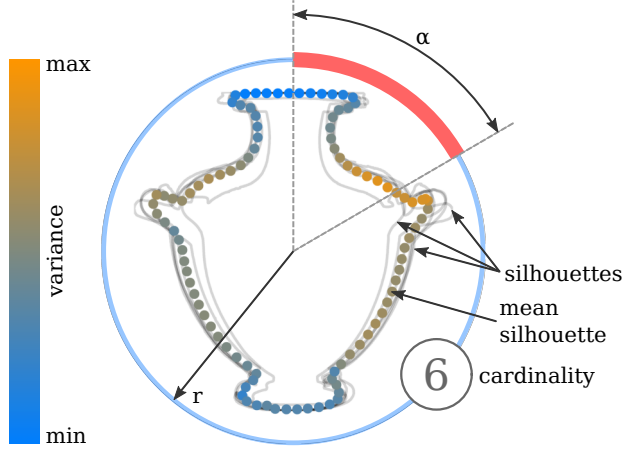


FIGURE 7.2. The SPS indicates the shapes of multiple aggregated objects by both their overlaid silhouettes as well as a mean silhouette (colored dots) computed over all surrogated objects. The color of the dots refers to the local shape variance. A circled number at the bottom right indicates the number of objects which are represented by the MOP, which is also hinted at by its size. The red arc provides a visual cue regarding how many of the preview’s objects are part of the current highlighting.

7.3.3.1 Hierarchic Balloons

The system also has to support the on-demand dynamic reversing of aggregations to allow users to examine the data at all levels of visual granularity as required by **T4**. By now, this kind of exploration is implemented for the SSV only, where we make use of the aggregated graph data structure $G^{SSV}(V_{agg}, E)$. Although (some) vertices serve as surrogates for others, they store a hierarchic list of vertices and edges. *I.e.*, the data structure at hand constitutes a tree of graphs with the leaves corresponding to single artifacts $G_{leaf}(\{.\}, \emptyset)$. A proper visualization technique for this kind of data structure are so-called *balloon views* [278]. The basic idea is to encapsulate subtrees in circles attached to the parent node, isolating it visually while still preserving its relation to the next higher level in the hierarchy. In our implementation, a balloon has its own force layout and is subjected to the same aggregation and sparsification operations as the graph of the root node, meaning that it can contain aggregated nodes as well. Those can also be turned into balloons recursively, down to the lowest level of the tree. Balloons are completely agnostic *w.r.t.* nodes higher up the hierarchy, and incorporated balloons are considered only by means of their radius, determined by the extent of their force layout, which serves as a parameter for the collision force. Balloons are added only on demand to avoid visual overloading. That is, a user

has to click its corresponding MOP (FIG. 7.3). Note that a balloon is filled with an arbitrarily selected color which is restricted to shades that have a sufficiently high contrast *w.r.t.* its parent's color. The system supports the concurrent display of multiple balloons from different levels of the hierarchy as well as from the same level.

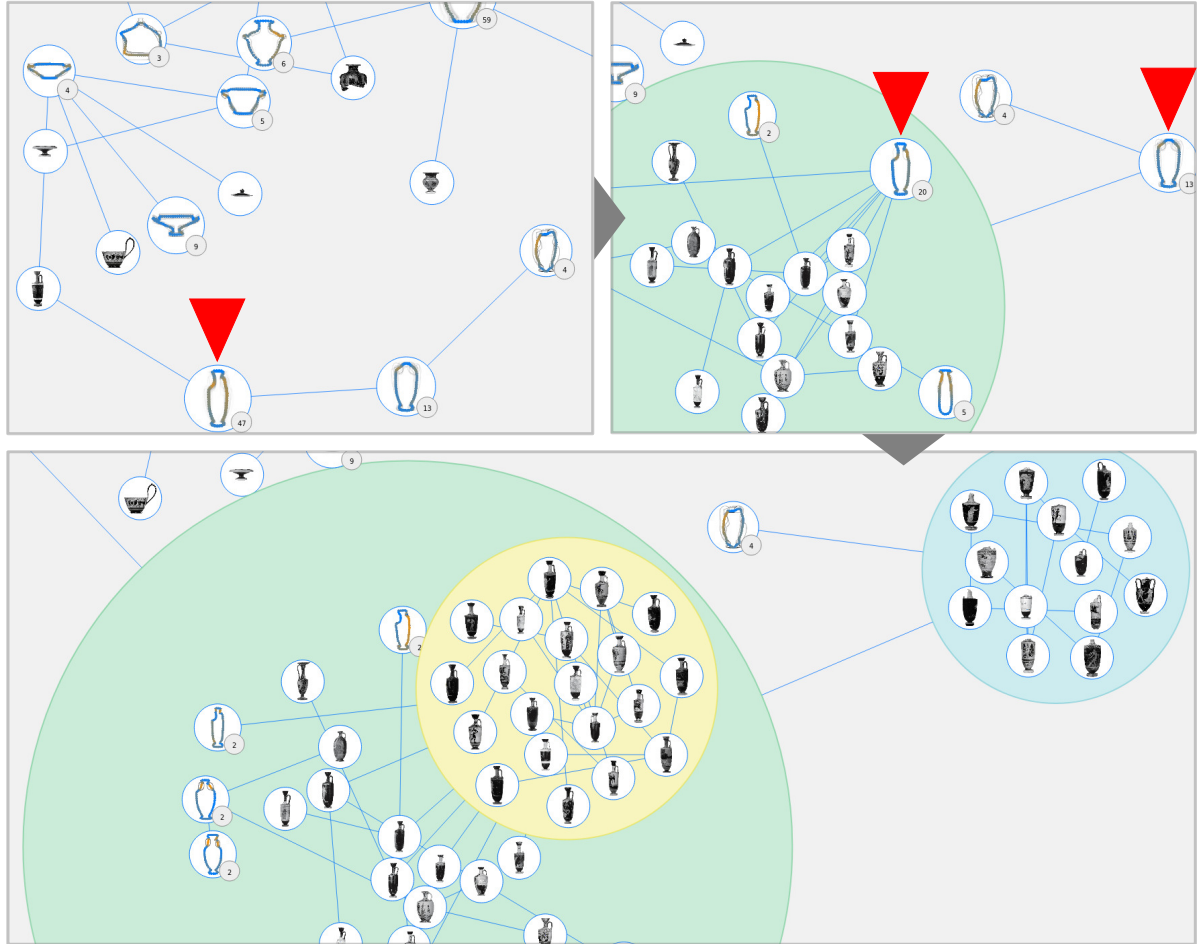


FIGURE 7.3. Balloon views support the recursive exploration of a SPS's corresponding objects. The top left figure shows, among others, a cluster of 47 lekythoi shapes. Clicking this MOP (click interactions symbolized with red triangles) opens the balloon view in the top right figures. Bottom: multiple balloons at different recursion depths can be displayed concurrently.

7.3.3.2 Intra-view Highlighting and Linking

The linking of views in a multiple linked views system is referred to by *highlighting*. For our purpose, a lasso selection turned out to be the easiest to use for selecting. For visualizing highlights, we decide on a change in color, as a preview's size already encodes cardinality information (SEC. 7.3.2.1). A challenge poses the different aggregations in different views, for oftentimes only a fraction of the artifacts related to a MOP belongs to the current highlight. We

visualize this partial highlight with a red arc around a MOP (FIG. 7.2), whose angle α amounts to $\alpha = |\#highlights'|/|\#artifacts'| \cdot 2\pi$. Due to the navigation capabilities of the views (panning and zooming), not all previews are always within the currently displayed section. If this is the case for a highlighted preview, we want to indicate their presence regardless and do so via *signposts*.

Still, this form of highlighting does not allow drawing conclusions regarding the correlation of clusters across separate views. This information is revealed using an alluvial-flow diagram, comparable to the diagram presented by Saunders [279], between all views. The streams start and end at the respective previews, and the stream thickness corresponds to the number of highlights it connects (FIG. 7.6). A stream is only drawn if both the source and the target preview are within the currently displayed sections.

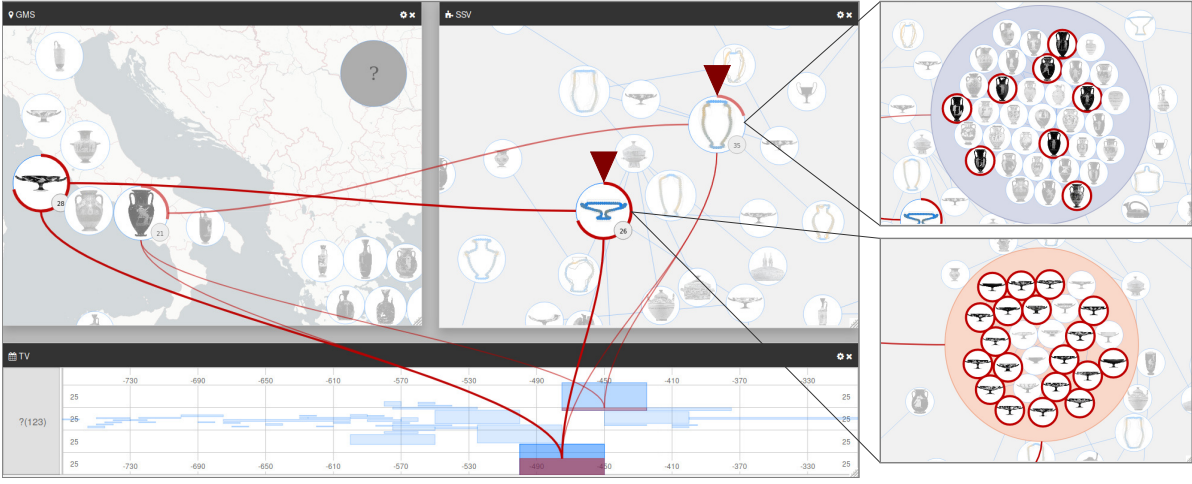


FIGURE 7.4. The LVVES with intra-view links based on the most significant relations revealed by the intra-view correlation. Nolan amphorae and cups preferred in Etruria are detected. The two balloon views on the right side correspond to the highlighted SPSs in the SSV.

7.3.3.3 Intra-view Correlation

Besides a user-defined selection of intra-view links, we also pre-calculate the ‘most significant’ selections. *I.e.*, we determine those artifacts whose neighborhoods are the most similar across all views. Since we have graph data structures with a common set of vertices across all our views (G^{GMV} , G^{TV} , and G^{SSV}) this is achieved through a vertex-wise correlation measure for different graph layouts. The intra-view correlation $(G^i \star G^j)(v)$ for a vertex $v \in V$, common across graphs $G^i(V, E^i)$ and $G^j(V, E^j)$, is thus proportional to the similarity of its neighborhoods with $(G^i \star G^j)(v) \propto \text{Sim}_E(E_v^i, E_v^j)$, with E_v^i, E_v^j as the respective subsets of edges incident to v . For n_G graphs and their set of edges $\mathcal{E} = \{E^i \mid i \in [0..n_G]\}$ the cross-correlation of v is given by

$$(7.4) \quad \text{Corr}(v, \mathcal{E}) = \frac{1}{\binom{n}{2}} \sum_{\langle E^0, E^1 \rangle \in \binom{\mathcal{E}}{2}} \text{Sim}_E(E_v^0, E_v^1).$$

For the SSV this measure is directly applicable as it has a complete graph $G^{SSV}(V, E^{SSV})$ defined by its adjacency matrix. The graphs of the TV and the GMV, on the other hand, do not have any edges $G^{GMV} = G^{TV} = G(V, \emptyset)$, resulting in a uniform cross-correlation of 0. This drawback is compensated by adding artificial edges between all vertices for G^{GMV} and G^{TV} . For the GMV, the weight of edges between vertices belonging to the same cluster (and thereby same location) is set to 0, while all others are initialized with a weight relative to their geographic distance. For the TV, the weight of an edge between vertices belonging to the same date range is set to 0 while the weight of an edge between two vertices $v_i, v_j \in V$, belonging to different date ranges is defined as the difference $d_{TV}(v_i, v_j) = |z(v_i) - z(v_j)|$, where $z(v)$ indicates the center of the date interval associated with vertex v . Note, that no normalization of edge weights is necessary as they are exclusively used to determine an ordering. The resulting intra-view correlations are sorted by value, and only the most prominent ones, down to a user-defined threshold, are displayed via the highlighting and parallel links systems (SEC. 7.3.3.2), with the strength of the correlation encoded into the transparency of the links and associated previews (FIG. 7.4).

7.3.4 Results and Evaluation

To evaluate our approach we implement a prototype as a web application that relies heavily on the D3 JavaScript library in connection with HTML5 Canvas. A C++ backend is responsible for the computation of the shape similarity required by the SSV (SEC. 7.3.2.3) as well as for the intra-view correlation (SEC. 7.3.3.3). The geocoding necessary for the GMV is obtained from the OpenCage Geocoder API¹.

The LVVES has been used by domain experts for the analysis of real-world archaeological data, following the different analysis tasks in SEC. 7.3.1. In terms of data, we need a set of objects having both pictorial data and metadata. From TAB. 2.2 in SEC. 2.1.2, we select the five CVA fascicules CVA Dresden 2; CVA Erlangen 2; CVA Göttingen 3; CVA Jena 1; and CVA Munich 16, whose records we combine with metadata obtained from the BAPD (SEC. 2.1.1). In total, this results in a database comprising 614 vases with 2,799 images. This initial selection is filtered automatically (based on metadata) to exclude fragments that do not exhibit meaningful shape information. From the remaining 317 artifacts, the associated images have been filtered manually (SEC. 7.3.2.3) by removing close-ups and sketches. 61% of these vases have dating information, while only 32% have associated provenance information. This drawback is typically for museum collections of antiquities which are mostly compiled of artifacts coming from the art market since the 18th cent. CE. [280].

7.3.4.1 Application and Domain Expert Evaluation

We select three specific showcases to evaluate the system's applicability for actual archeological research and demonstrate some results. For each of them, we use a different combination of

¹<https://opencagedata.com/>

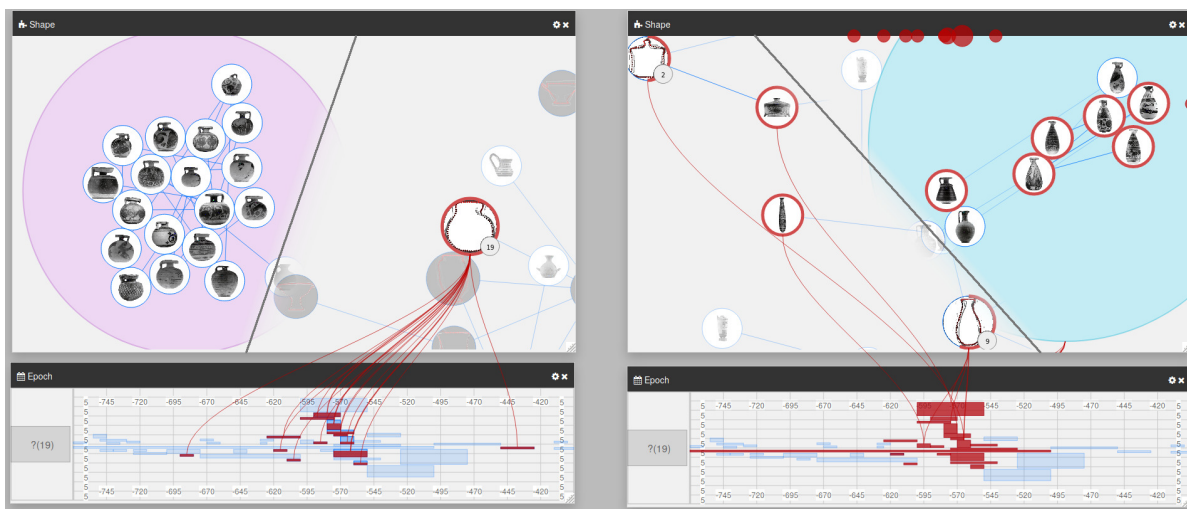


FIGURE 7.5. Left: A group of aryballois is selected via their SPS in the SSV to examine their different datings. Right: A selection of the strongly represented date range in the TV unveils which other shapes appear during this epoch. *E.g.*, a group of 9 alabastra. The balloon views illustrate which artifacts are surrogated by the SPSs.

modalities (date, provenance, shape). For **T1**, we use the GMV to investigate artifacts' provenances; for **T2**, the TV to analyze the temporal distributions of artifacts and for **T3**, the SSV to get an overview of shape variations.

For the first showcase, related to **T3**, **T4**, we take the data sets of Erlangen, Göttingen, and Dresden and look at their representation in the SSV. Greek pottery was produced in different shapes and sizes according to the intended use (storage, drinking, containing perfume, *etc.*). Although the main shape types of Greek pottery remained relatively constant due to their function, there were a lot of varieties and differences in the accentuation of details [15]. *E.g.*, one specific shape, the lekythos, a slender, single-handled, and narrow-necked vessel, was used for storing oil. The SSV displays this common shape of lekythoi efficiently with a SPS of 47 artifacts (FIG. 7.3 top left). The SPS allows to us to recognize at a glance the low variability of the individual shape profiles of these clustered artifacts, as only small differences can be seen in the transition from the shoulder to the neck of these vessels. The dynamic level of visual granularity from overview to close-up by means of hierarchic balloons (SEC. 7.3.3.1) facilitates the examination of linked vases in detail. The first balloon (FIG. 7.3 top right) shows some varieties of lekythoi and a further extensive SPS of 20 artifacts. The second balloon (pink) demonstrates that SSV has clustered herein the lekythoi of the canonical proportions with a narrow, almost cylindrical body. A limitation pose degenerated vases, *e.g.*, different vase types with broken mouths are joined in one SPS in FIG. 7.3 bottom.

For the second showcase, addressing **T5**, we look at the vase collection of the university Jena in its entirety, more specifically on their represented shapes in relation to their given date ranges.

Within the largest cluster, depicted in FIG. 7.5 left, the SSV grouped together vessels with a globular body. Apart from two exceptions – an early Protocorinthian aryballos and a late belly lekythos – all exhibit a shape known as *round aryballos*, a perfume vessel with a spherical body, a narrow neck, and a broad disc-like mouth, which becomes regular at the end of the third quarter of the 7th cent. BCE. in Corinth. This shape underwent mass production with a wide distribution in the first half of the 6th cent. BCE and disappears almost completely in the Greek sphere after the mid-6th cent. BCE. The balloon in FIG. 7.5 left also shows imitations of this specific Corinthian shape in other vase production centers: East Greek/Rhodes (made of Faience) and Italy (the so-called Corinthianizing pottery). Suppose we investigate which other shapes exist from the first half of the 6th cent. BCE. At Jena, the SSV offers, among some single shape types, a second larger group (FIG. 7.5 right). In this group, all Corinthian and Corinthianizing alabastra at Jena were grouped together (with two outliers). To sum it up, the LVVES clearly supports the investigation of typo-chronological developments of vessel shapes and can enable the recognition of new strands of developments or peculiarities with larger datasets.

The third showcase is a characteristic example for **T6**, connecting shape, provenance, and date (FIG. 7.4). Neck amphorae in a distinctive slim version and with an elongated neck are clustered by the SSV while the GMV reveals that the vast majority of this type was found in Campania, Southern Italy, especially in the ancient city of Nola. This type of amphorae was named, following the archaeological site, *Nolan amphorae*. Even though this is not a new discovery, the visualization emphasizes this export from Athens to Southern Italy. The ancient region of Etruria in Central Italy is a preferred exportation site for another vase shape, the cup. The timeline outlines that these exports started later and lasted longer than the exports of Nolan amphorae to Southern Italy. Looking at the pottery market in general, it is an interesting archaeological fact that specific shapes were popular in different regions of Italy.

7.3.5 Discussion Limitations and Open Research Questions

An open question is how to determine a sensible amount and routing for the parallel links in order to achieve the optimal balance between a maximum of visible information and a sensory overload, as is the case in FIG. 7.6. This problem could potentially be addressed with tailored filtering approaches as well as methods routing with reduced collisions like the method presented by Steinberger et al. [281].

Another challenge are historic toponyms in artifacts' provenances and other spatial metadata which have no equivalent in modern geocoding services. While we used OpenCage Geocoder for the sake of convenience, many ancient city names cannot be resolved by this service. If the demonstrator is further developed for everyday professional use, it might be worth utilizing a gazetteer which is specialized in ancient place names like the Pleiades project², which was unfortunately not yet accessible at the time the experiments were conducted.

²<https://pleiades.stoa.org/home>

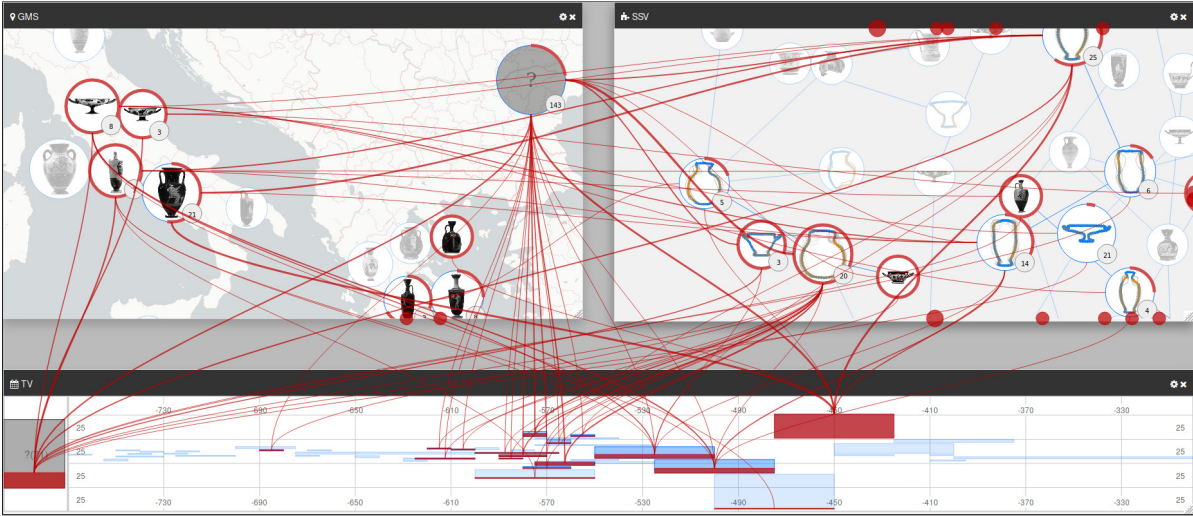


FIGURE 7.6. Too many parallel links at the same time congest the LVVES and can hinder the detection of otherwise visible correlations.

We want to note, that even though the modalities supported by the prototype implementation are very limited (date, provenance, and shape), the proposed views are, to an extent, independent from the visualized modality and can be readily employed for other object traits. *I.e.*, the GMV can visualize any spatial data (*e.g.*, production sites, current collections) while the SSV can display any non-categorical data (*e.g.*, painting style) where relations between artifacts can be specified by a distance function.

7.4 Conclusion

Gaining insight and understanding of CH artifacts across modal boundaries is an archaeologist's day-to-day work. Available data in the form of 3D scans, images, and a variety of metadata helps to analyze not only the economic situation of producers and recipients but the ancient social systems in general which is a core research question of Classical Studies. The proposed visual exploration system (LVVES) addresses the need for an integrated exploration system for a selection of modalities, which it brings together using visual links. Those can be triggered either by interaction or automatically based on a well-defined correlation measure. First experiments on a small-sized test dataset proved that the latter is able to reveal object correlations that are known to be present in the data. Even though we do not claim that this system is able to resolve all challenges related to the problem at hand, we believe that this (to our knowledge) novel design poses a valuable contribution to the field of visualization of CH data and offers great potential for scalability and extendability.

Part IV

Digital Restoration

COLORIMETRIC RESTORATION

Contents

8.1	Introduction	161
8.2	Related Work	163
8.3	Content-based Repetitive Pattern Completion Workflow	166
8.3.1	Generation Rule Detection	167
8.3.2	Content-based Reconstruction	169
8.3.3	Results	171
8.3.4	Discussion, Limitations and Future Work	175
8.3.5	Applications	177
8.4	Conclusion	178

Digital restoration constitutes the smallest thematic part of this thesis and was originally no intended research objective within the CrossSAVE-CH project. It was, however, discovered, while researching analysis methods for repetitive surface patterns, that the very structured arrangements of such patterns can be leveraged for both detecting gaps in the arrangement as well as for generating texture hypotheses for these areas if the majority of a pattern sequence is preserved. That is, we focus exclusively on the colorimetric restoration of artifacts, something that has not yet been heavily researched for 3D pottery or, to our knowledge, for any 3D data whatsoever.

8.1 Introduction

The digital restoration of CH artifacts is a well know research objective, as the physical restoration is usually not possible due to various reasons, *e.g.*, cost factors, the fragility of objects, or different parts of one and the same object being curated by different museums. Also, the advancement of

This chapter is based on:

- Lengauer, S., Preiner, R., Sipiran, I., Karl, S., Trinkl, E., Bustos, B., Schreck, T., “Context-based Surface Pattern Completion of Ancient Pottery”, in *Eurographics Workshop on Graphics and Cultural Heritage*, The Eurographics Association, 2022. DOI: [10.2312/gch.20221234](https://doi.org/10.2312/gch.20221234)

3D scanning and the increasing availability of detailed 3D models encourage the development and application of computer-based restoration techniques. Papaioannou *et al.* [282] differentiate thereby between *reassembly*, the identification of affiliations between fractured artifact parts and their realignment, and *object completion*, the approximation of object parts that are missing due to erosion or an incomplete set of object fragments. Our proposed approach belongs to the latter, but in contrast to the majority of techniques that deal with the completion of an object’s geometry [22–28, 283–286], we focus exclusively on the completion of its texture. More specifically, we focus on the completion of repetitive patterns which exhibit at least some degree of structured arrangement because, in contrast to painting style, color scheme, and illustrations, which vary greatly between different cultures and periods, sequences of repetitive patterns are an almost ubiquitous theme. They are usually arranged in rows or columns around the objects’ solid of revolution. Due to the various stages of degradation characteristic of ancient pottery, elements of a sequence are oftentimes eroded beyond recognition or missing altogether, as can be seen in the examples given in FIG. 8.1. Following the geometric completion concepts, which leverage rotational symmetry aspects, we exploit concepts of regularity and similarity in order to reconstruct missing or eroded parts. With our proposed approach, we aim to complete missing patterns in incomplete ornamentation sequences. While the completion of missing image parts is a well-known image processing task, known as *inpainting* [205, 287–289] (*c.f.* query inpainting for 3D+Sketch in SEC. 4.4.2.2), such approaches are generally agnostic *w.r.t.* global pattern



FIGURE 8.1. A selection of pottery from the MAJRCexhibiting various stages of erosion of their surfaces. The respective areas are highlighted with red polylines.

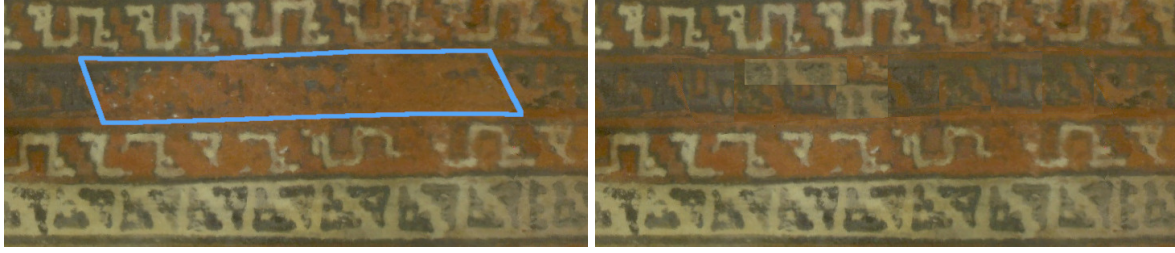


FIGURE 8.2. Left: The original surface of the vessel depicted in FIG. 8.1, right. The blue polygon marks the region to be inpainted. Right: The result after applying a patch-inpainting approach [205].

arrangements and are not aware of any missing patterns given an empty space on the surface to fill. In FIG. 8.2, we provide an example of the outcome of the patch-inpainting approach by Criminisi *et al.* [205], applied to the surface of the pre-Columbian pot shown in FIG. 8.1, right. It can be seen that, while the generated texture at the vicinity of the mask borders appears to be indistinguishable from the original texture, the result gradually worsens towards the center of the inpainted region where the generated texture is a mix of several unrelated image regions. This is to be expected as general-purpose inpainting algorithms are not able to capture the essence of the sequence but merely reconstruct the target area based on the border gradients and colors. We conclude that of-the-shelf inpainting algorithms are generally not able to faithfully reconstruct ornamentation sequences.

In the following, we present a workflow (SEC. 8.3) where we estimate the positions, orientations, and scales of missing patterns based on the sequence’s generation grammar before we employ a statistical denoising of the respective areas based on the remaining elements of the sequence. Before that, we will take a look at related work from this area (SEC. 8.2). Finally, we conclude the chapter with a summarization of our thoughts regarding the applicability of restoration and a discussion of open research questions (SEC. 8.4).

8.2 Related Work

As mentioned above, approaches for the digital restoration of ceramics can be categorized into approaches for geometric and approaches for colorimetric restoration.

Geometric Restoration. Restoration of defects in digital 3D cultural heritage objects has several aspects and requires solving different tasks. That is, it has been an intensive research focus, especially over the last decade. The first attempts to tackle the problem of reproducing surface patterns in CH objects were mainly focused on geometric methods involving user intervention. Kolomenkin *et al.* [290] proposed a framework to recreate surface relief patterns from line drawings entered by a user. The method optimizes the 3D relief geometry for a given drawing by imposing smoothness constraints characterized by the Laplacian of the relief. These

constraints enhance the smooth transitions between the reliefs and give a realistic appearance to the reconstructions.

Andreadis *et al.* [23] propose a semi-automatic pipeline for the geometric reassembly of fractured 3D objects. The pipeline consists of several steps. First, a preprocessing step produces a segmentation of the object surfaces and classifies each segment as either ‘intact’ or ‘fractured’. Second, the method matches the fragments according to their fractured surfaces. Third, the method performs a pairwise surface registration for fragments exhibiting significant erosion or a missing part. Fourth, the object is reassembled using a graph-based approach. Finally, in the last step, symmetry-based fragment registration is used to align any remaining disconnected fragment. Each of these steps requires human interaction to ensure a correct reassembly process and to help the system align fragments that are too eroded or have significant missing parts.

Lamb *et al.* [286] propose an automatic method to reconstruct the missing geometry of a 3D object. To this end, the algorithm requires a reference 3D object. In the first step, the incomplete object is aligned with this reference object. The exterior of the missing part is generated using the information from the alignment step. Finally, the algorithm computes a smooth surface transition along the outlines of the fracture. The generated fragment can be 3D printed for restoration of the broken object. Setty and Mudenagudi [291] propose a similar reconstruction method of missing parts for 3D point clouds. The method relies on a set of ‘exemplar’ models, which are geometrically similar objects to the one that is being repaired. The method selects the regions from the exemplar objects that best fit the missing parts and uses them for the reconstruction of the object.

In contrast to these exemplar-based methods [286, 291], ‘hole filling’ approaches [26–28] try to reproduce a missing part of an object’s geometry by patching it with a context-based or coherent surface patch obtained from the input object itself. Sharf *et al.* [26] propose an automatic method to fill holes or to complete point-sampled surface areas in 3D models. The main idea of this method is to fill the hole with a surface patch considering its context, that is, the characteristics of the surface surrounding the hole. The authors show some examples of the results obtained with their method and compare it to patching the hole with a smooth surface. They also show that their method is robust to noise. A disadvantage of this method is that no overall context is considered, but holes are just filled such that the patch is merely ‘similar’ to its surroundings. Harary *et al.* [27] propose a method for the same problem, with the difference being that it is not only context-based but also ‘coherent’. The coherence restriction imposes that the local patch must be similar to some local patch in the model. This helps avoid some errors that can be produced by context-based methods, where the patched surface may be similar to its surroundings but significantly different from a corresponding local neighborhood in the surface (for example, a symmetric counterpart of the patched surface). In addition, orthogonally to that, the method by Harary *et al.* [28] follows a semi-automatic concept, and requires the user to input four points that are used to compute a curve. This curve defines ‘triangle strips’, which are considered

constraints of the problem and aid the reconstruction process. If the surface hole covers two different segments or parts of the 3D model (*e.g.*, if it covers part of the hair and part of an ear of a bust), and the user-defined strips separate these segments, the method can produce coherent and context-based patches for each of these segments.

Papaioannou *et al.* [282] propose a complete geometry restoration workflow for diverse CH objects comprised of reassembly and completion steps. For the prior, they leverage both fracture surfaces as well as salient features. For the latter, they present a procedure based on generalized symmetries and a complementary part extraction. They demonstrate their pipeline on a number of varied CH objects. The publication is part of the European Union-funded PREdictive digitization, reStoration and degradatIOOn assessment of CH objectS (PRESIOUS) project¹.

For wheel-made pottery, the largest group of CH objects, symmetry properties can be leveraged for completion efforts. Sipiran [24] describes a method that completes missing geometry by replicating symmetric features. In particular, this method exploits the global-local behavior of a heat-based function defined on the 3D surface. The function is formally proven to be invariant to rotational symmetry and exhibits symmetric correspondences around the generator axis of a given object. The method is able to detect the symmetry axis of incomplete solids of revolution. This information is then used to replicate missing parts. In a later publication, Sipiran [25] explores the suitability of fast local feature detectors to detect repeatable regions along the rim of rotationally symmetric objects. Since these regions convey enough information about the symmetric structure of an object, they can be used to replicate existing geometry for completion purposes. The author also demonstrates the effectiveness of the method for CH objects.

Colorimetric Restoration. The problem of colorimetric restoration is orthogonal to the geometric one and usually relies on image restoration techniques to this end. In contrast to geometric restoration, it is thematically broader. Several previous efforts deal with the related problem of the colorimetric restoration of paintings and mosaics [292, 293]. Yet, the colorimetric restoration of the 3D surfaces of other CH objects has not received much attention until recently.

One of the first works was done by Aliaga and Law [30], who present a concept for both the colorimetric restoration as well as presentation of pottery objects. The restoration relies on the input from a user who specifies a set of relevant colors together with example pixels for each color from all over the object surface, which is given as a photograph. This input is used for a sophisticated labeling technique that segments the surface into areas of coherent colorization. The user can influence this stage through sketching and brushing tools. Finally, the outcome of the restoration is projected on the original object allowing multiple viewers to observe the restored object without the need for any viewing devices.

Later, Gilad-Glickman and Shimshoni [31] propose a method denoising 2D as well as 3D CH artifacts. To this end, they assume that decorations appear in a repetitive manner and with a

¹<http://www.presious.eu>

limited palette of colors. The basic idea is to determine cohesive regions based on color, using a Graph-Cut segmentation (*c.f.* SEC. 3.2.1). The underlying graph is enhanced with additional edges introducing a semantic context by connecting pixels having a similar neighborhood. Hence, these edges should connect areas depicting the same letters or ornament and serve as a cue for similar colorization. The applicability of the approach is demonstrated on both 2D as well as 3D inputs.

One of the most recent works on colorimetric completion within a CH context is given by Assael *et al.* [29], who employ a deep-learning architecture based on transformers with the aim of restoring ancient Greek inscriptions. The proposed deep neural network, dubbed *Ithaca*, can predict missing characters in these ancient texts. It can also predict the geographical origin of a text among 84 different regions and its period of creation in a time span from 800 BCE to 800 CE. To this end, the network was specifically trained with data from texts in the ancient Greek language.

Positioning of Our Work. It appears the majority of related work is focused on geometric reconstruction, while publications dealing with colorimetric completion of CH objects are rather sparse and mostly limited to approaches for image restoration. For this task of image-inpainting, an abundance of methods was proposed over the years [294]. The most relevant prior works for our concept are given by Aliaga and Law [30] and Gilad-Glickman and Shimshoni [31]. However, the prior relies purely on user input in order to recognize and outline missing patterns, while the latter also requires largely intact patterns for the denoising to work properly. There is to our knowledge, no approach that is able to determine the absence of patterns or complete partially preserved patterns. Also, both of the mentioned approaches assume a very limited number of colors and therefore segment the surface into regions of equal color. In contrast to this rather artistic representation, our approach leans more towards a visually realistic restoration with color variations and gradients that can also be observed within coherent surface regions.

8.3 Content-based Repetitive Pattern Completion Workflow

FIG. 8.3 illustrates our proposed reconstruction pipeline. The input, as well as the output, are 3D models which differ only in their texture. Similar to Gilad-Glickman and Shimshoni [31], we reduce the complexity of the colorimetric restoration by mapping the 3D surface into 2D (*c.f.* SEC. 3.1). *I.e.*, all processing in-between is conducted in image space. As a first step all entities of a given pattern class, together with their scale and orientation, are determined in a *Self-Similarity Detection* step. From this set of unordered pattern entities, a sequence is determined by leveraging simple heuristics intrinsic to pottery objects. Thereafter, the transformations between consecutive entities are computed and used to formulate the sequence's generation grammar. All these steps are described in depth in SEC. 8.3.1. If the estimated transformation between two neighbors differs greatly from the expected approximated prototypic transformation,



FIGURE 8.3. Pattern completion pipeline. The input model surface is unwrapped and subject to a self-similarity detection of individual pattern entities belonging to different classes. Individual sequences of repetitive entities within these sets are then detected, and their generative grammar is estimated. These are inputs to a context-based reconstruction and substitution of missing or deteriorated pattern elements in gaps detected in these sequences (green bracket). The reconstructed texture is finally mapped back onto the input geometry. Here, two ornament sequences of snake patterns are completed (red bracket).

it is taken as an indication of a gap in the sequence (SEC. 8.3.2.1). After determining the locations and transformations of reconstruction candidates, a statistical model is applied to complete the remaining texture in an optimal manner based on the other patterns of this class (SEC. 8.3.2.2).

In SEC. 8.3.3, we present the application of this workflow on a selection of real-world CH objects from pre-Columbian and Greek cultures. As an additional validation, we also experiment with a synthetic surface abrasion for objects which do not exhibit surface deficiencies. This has the advantage that we can increase the number of valid reconstruction targets while also having a ground truth for the completion process. Finally, we also state limitations and possible targets for future work (SEC. 8.3.4) before we conclude the section with a discussion of possible applications in the CH domain in SEC. 8.3.5.

8.3.1 Generation Rule Detection

Given an object with repetitive surface patterns, the appearance of a specific pattern entity within a sequence can be generally estimated using a generation rule. That is, we assume a Markovian process [295], meaning that the location, orientation, and size of a pattern entity can be fully described as a transformation applied to its predecessor. Within a sequence, these transformations between consecutive patterns are quite similar unless there is a gap. Hence, the first objective is to find the prototypic transformation, which describes the transformations between the uninterrupted parts of the sequence in an optimal manner. To this end, we use a four-stage process involving: (i) the detection of the individual occurrences of a pattern, (ii) the detection of their inherent sequence (if present), (iii) the determination of the transformations between neighbors, and (iv) the estimation of the prototypic transformation. Prior to the first stage, we project the input, a textured 3D model, to a 2D plane. This greatly simplifies all processing

steps and is possible for most rotationally symmetric shapes without introducing much distortion. At the end of the pipeline, the inverse projection is used to remap the reconstructed texture back onto the 3D model. Several methods for the unwrapping of 3D surfaces have been published. Specifically, for the CH domain, approaches using proxy geometries [153] or mass-spring systems [77] have been proposed. In our experiments, we use a variation of the cylindrical unwrapping by Karras *et al.* [76].

The detection of the individual entities of a given surface pattern can be formulated as a self-similarity detection problem [81, 296–299], a very well-researched topic in computer vision. Note that this part of our pipeline can be performed by any of these methods providing areas and orientations of target entities and is thus not within the scope of this paper. In our experiments, we rely on the annotated pottery benchmark dataset (SEC. 8.3.3). That is, we assume as input a set of unordered pattern entities $\{S_i\}_{i \in I}$, with the index set I . Each entity has to convey (at least) the attributes orientation and scale (*w.r.t.* a reference/pivot entity) as well as location. Yet, scale and location properties can also be derived if the entities are given as areas (sets of pixels or surface parts). Based on this input we determine the presence of structured arrangements. More specifically, in the case of ancient pottery, repetitive patterns are commonly arranged in rows or – in rare cases – also in columns. Patterns that are scattered in a random fashion all over the object’s surface are beyond the scope of our approach. We estimate the locations of rows K at offsets $Y = \{y\}_{k \in K} \subset \mathbb{R}_+$ and columns L at offsets $X = \{x\}_{l \in L} \subset \mathbb{R}_+$ based on the pattern entities’ centroid locations $\mathbf{C} = \{C_i\}_{i \in I} \subset \mathbb{R}^2$. To this end, a GMM is fitted to their respective x and y -coordinates. The most likely number of mixtures, reflecting the number of rows or columns, is determined using the Bayesian Information Criterion (BIC) [300] by comparing the quality of the fits with $n \in [1..n_{max}]$ components. The maximum number of rows or columns n_{max} is limited to $|I|/3$ as we require each row or column to feature at least 3 pattern entities. I_k and I_l denote the index sets of pattern entities belonging to the k -th row (sorted by their x -coordinate) and the l -th column (sorted by their y -coordinate), respectively.

The relation between the pattern S_i and its predecessor S_{i-1} can be formulated as $S_i = \hat{M}_{i-1,i} S_{i-1}$, with the most likely transformation $\hat{M}_{i-1,i} = \operatorname{argmax}_{M \in \mathcal{M}} \operatorname{sim}(S_i, M S_{i-1})$ and an arbitrary similarity function sim . The task of finding $\hat{M}_{i-1,i}$ in the set of possible transformations $\mathcal{M}$ is an image registration problem [301]. For our application, we assume that $\hat{M}_{i-1,i}$ can be sufficiently well described with an affine transformation. With this class of geometric transformations, a combination of a linear transformation and a translation can be modeled. In the 2D domain, this transformation is usually represented with a 2×3 matrix

$$(8.1) \quad M = \begin{bmatrix} c_x \cos(\phi) & -s_x \sin(\phi) & t_x \\ s_y \sin(\phi) & c_y \cos(\phi) & t_y \end{bmatrix},$$

with c_x and c_y as the scaling, s_x and s_y as the shear, ϕ as the rotation and t_x and t_y as the translation. For our application, we assume a uniform scaling $c = c_x = c_y$ and the absence of any shear $s_x = s_y = 0$. We denote the transformation from pattern $i - 1$ to pattern i with

four parameters as $\theta_i = \langle c_i, \phi_i, t_{x_i}, t_{y_i} \rangle^T$. To determine the prototypic transformation between any consecutive patterns, with parameters $\mu = \langle \tilde{c}, \tilde{\phi}, \tilde{t}_x, \tilde{t}_y \rangle^T$, we fit a multivariate *Students t-distribution* [302] $\text{St}(\Theta|\mu, \lambda, \nu)$ to the set of transformations $\Theta = \{\theta_i\}_{i \in I}$, with ν as the degrees for freedom and λ as the inverse precision scaling. We selected this distribution model as it is robust *w.r.t.* outliers for a small sample size.

8.3.2 Content-based Reconstruction

The content-based pattern recognition is a two-stage process. First, gaps in the pattern sequence are detected (SEC. 8.3.2.1); before the number of missing patterns, their locations, orientations, and visual appearance is approximated in a gap filling step (SEC. 8.3.2.2).

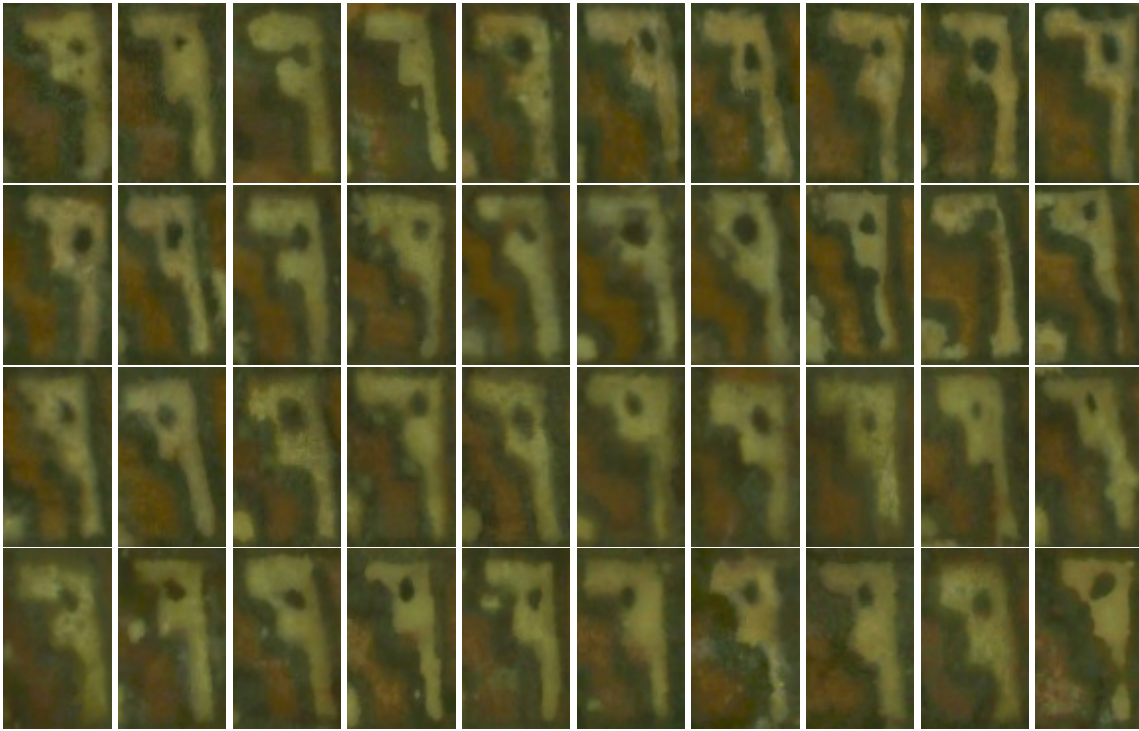


FIGURE 8.4. A selection of 40 reference samples of a staircase pattern, exhibited by *MAJRC 0185*, after alignment and histogram matching.

8.3.2.1 Gap Detection

For the identification of gaps, we leverage the prototypic transformation parameters μ . We define gap candidates as the patterns whose transformation *w.r.t.* their predecessor, given by θ_i , differs significantly from μ . That is, the index set of gap candidates $\mathcal{G}$ is formally given by

$$(8.2) \quad \mathcal{G} = \{i \in I \mid \max|\theta_i - \mu| > t_g \sigma\},$$

with $\sigma = \sqrt{|I|^{-1} \sum_{i \in I} (\theta_i - \mu)^2}$ as the standard deviations of the transformation parameters and t_g as a cut-off threshold. For practical applications, $t_g = 2$ showed to provide satisfying results. Note that erroneous gap candidates, resulting from a too-low threshold, are usually no concern as it will be recognized in the gap filling step that no pattern will fit in-between.

8.3.2.2 Gap Filling

For each of the gap candidates $g \in \mathcal{G}$ (Eq. (8.2)) we add $n_g = \max\{\theta_g \odot \mu\} - 1$ ($\odot$ denotes the element-wise Hadamard-division) patterns to fill the void. We determine their respective transformations such that the space between the last pattern before the gap and the next preserved pattern is optimally bridged. *I.e.*, we expect the bounding box $bbox_j$ of the j -th new pattern to be at $bbox(S_{g-1})M(\theta_g \odot j + 1)$, with $M(\theta)$ as the affine transformation matrix (Eq. (8.1)) for θ .

In order to estimate the original texture at $bbox_j$ we perform Bayesian image denoising [303], which allows us to optimally infer the noise-free version $\mathbf{y}$ of a noisy input $\mathbf{x} = [x_0, x_1, \dots, x_{D-1}]^T$, with $D = w \cdot h$ and w, h as the image's width and height. With $\mathbf{y}^* = [y_0^*, y_1^*, \dots, y_{D-1}^*]^T$ denoting the corresponding noise-free image, we formulate the relation between $\mathbf{x}$ and $\mathbf{y}^*$ as $(\mathbf{x})_i = (\mathbf{y}^*)_i \xi$, $i \in [0 \dots D)$, with ξ being sampled from a Gamma distribution, as we assume the noise in $\mathbf{y}$ to be multiplicative and independent. As $\mathbf{y}^*$ is unknown, the goal is to find the best-fitting $\hat{\mathbf{y}}$, given the observed $\mathbf{x}$. Such can be achieved with the Minimum Mean Squared Error (MMSE) function

$$\hat{\mathbf{y}}_{MMSE}(\mathbf{x}) = \frac{\sum_{\mathbf{y}' \in \mathbf{Y}} \mathbf{y}' p(\mathbf{x}|\mathbf{y}')}{\sum_{\mathbf{y}' \in \mathbf{Y}} p(\mathbf{x}|\mathbf{y}')},$$

with $\mathbf{Y}$ as the remaining noise-free entities of the pattern class and

$$p(\mathbf{x}|\mathbf{y}') = \frac{1}{\eta^{kD} \Gamma(k)^D} \exp\left(-\frac{\lambda}{\eta} \mathbf{1}^T (\mathbf{x} \odot \mathbf{y}') + \lambda(k-1) \mathbf{1}^T \log(\mathbf{x} \odot \mathbf{y}')\right)$$

as likelihood of the multiplicative Gamma noise. k and η are the Gamma distribution's shape and scale, respectively, while $\lambda \in [0, 1]$ is a governing parameter for the 'peakiness' of the likelihood. $\lambda = 1$ results in $\mathbf{x}$ being equivalent to exactly one element of $\mathbf{Y}$, while $\lambda < 1$ results in a blend from multiple of the best matching noise free images. $\lambda = 1e-2$ was used in all experiments.

In our case, the noisy input $\mathbf{y}$ is the texture observed at $bbox_j$ and the noise-free references $\mathbf{Y}$ are the textures at $\{bbox(S_i)\}_{i \in I}$. In order to employ the statistical model all inputs must be brought into a normalized representation, meaning that all images must have the same dimensions and display the pattern entities in the same manner. To this end, we, first of all, transform all elements to match the appearance of the pivot element. *I.e.*, for the pattern S_i , we apply the inverted transformation of all its prior sequence elements $\prod_{k=1}^{i-1} \hat{M}_{k-1,k}^{-1}$. Note that the transformations have to be applied in reverse order since affine transformations are not commutative. This aligns the bounding boxes of all patterns with the bounding box of the pivot element. Additionally, they must be resized to a uniform size of $\bar{w} \times \bar{h}$ pixels with $\bar{w} = |I|^{-1} \sum_{i \in I} width(bbox(S_i))$, $\bar{h} = |I|^{-1} \sum_{i \in I} height(bbox(S_i))$ as the patterns' mean width and

height respectively. Lastly, we match the patterns’ histograms with a sort-matching approach [304]. This algorithm allows matching the histogram of a source image to the histogram of a target image by sorting the pixels in both images by intensity. The source image is then assigned the pixel intensities of the target image according to the sorted sequence. We adapt this algorithm, originally intended for gray images, for colored inputs by performing the sorting and assignment for each channel separately in the *CIE L*a*b** color space [305]. The usage of this color space is justified by physiological studies which show that it is an appropriate representation of the color perception of the human eye. As we do not wish, however, to match images pairwise but bring a set of images into a common representation, we match each of them to the mean histogram computed over all images. A sample of the normalized reference images of a staircase pattern, exhibited by the pre-Columbian pot *MAJRC 0185* (FIG. 8.7), is featured in FIG. 8.4. The same normalization steps are also applied to the target texture at $bbox_j$. After obtaining the denoised version of the j -th new pattern, $\hat{y}_j$, we retransform it back to the predicted location and merge it with the original texture using a multiresolution spline technique [306].

8.3.3 Results

The following results section is tripartite. In the first part (SEC. 8.3.3.1), we visually assess the results obtained with our pattern completion approach on actual eroded real-world artifacts. In the second part (SEC. 8.3.3.3), we look at artificially degenerated surface paintings and compare the results of our pipeline to the original surface. To this end, we employ a custom-made *synthetic surface abrasion* process (SEC. 8.3.3.2), which aims to mimic the traits of actual wear. Finally, we also provide quantitative results (SEC. 8.3.3.4) by generating a large quantity of said synthetic abrasions and evaluate the performance of our method in the face of different degrees of degeneration by means of a quality measure.

8.3.3.1 Qualitative Assessment

For the qualitative assessment of our method, we look at two real-world objects exhibiting characteristic weather-worn deficiencies in their surface paintings. First, the pre-Columbian pot *MAJRC 0181* (FIG. 8.5, top) from the MAJRC. The annotation of its surface patterns results from a previous annotation effort (*c.f.* SEC. 6.3). This vessel features polychromatic surface paintings with two distinct pattern classes – a series of black triangular staircase patterns with an alternating orientation at the top and the bottom, as well as a snake-shaped pattern appearing in three parallel rows with black and white colorization. The pattern sequence in the middle row comprises four severely damaged pattern entities. From the results of the automatic completion (FIG. 8.5, top right), it can be seen that the newly generated texture blends in seamlessly and is almost indistinguishable from the original surface.

For the second exhibit – the amphora *KHM IV3600*– we annotate the repetitive patterns in a similar fashion to the pre-Columbian objects. Noteworthy *w.r.t.* pattern completion is the

sequence of linked bud and flower shapes on the lower body of the amphora, which is damaged in two areas (FIG. 8.5, bottom left). Yet, compared to the previous example, those are better preserved, and in some cases, there is just a fraction of the pattern missing. FIG. 8.5, bottom right, shows the results of the completion effort, which is used to reconstruct two pattern entities in the first gap and four in the second one. This example shows the benefits of Bayesian denoising, since the reconstruction is very faithful to the orientation and overall appearance of the incomplete patterns.



FIGURE 8.5. Left: The eroded original of *MAJRC 0181*, exhibiting a snake shape in three ornament bands together with its reconstruction. Right: The approach is also applicable to Attic pottery. In the example, a bud/flower shape on the lower body of the black-figured amphora *KHM IV3600* is reconstructed.

8.3.3.2 Synthetic Surface Abrasion

The examples in Section 8.3.3.1 are comprised of objects with real surface deficiencies, showing the applicability of our method for real-world data. On the downside, however, these do not allow us to objectively evaluate the quality of the completion since the original appearance of the surface pattern is unknown. As a consequence, we generate additional realistic examples by synthetically introducing plausible chippings to the object surfaces. The unchanged originals serve as the ground truth for the missing surface parts.

The synthetic holes are generated in a completely randomized manner on the unrolled surface, but with constraints that ensure that they resemble realistic chippings. That is, we model an arbitrary abrasion mask by starting with a circle with a random radius $r \in [r_{min}, r_{max}]$ ($r_{min} = 0.01d_s$ and $r_{max} = 0.05d_s$, with d_s as the diameter of the unrolled surface), which we place at a random location on the surface. To mimic the irregular shape and jaggedness of real holes, we superimpose a random noise ε to the radius of the circle, amounting to $\varepsilon(\theta) = \sum_{e \in [0..E]} a_e / e \sin(2^e \theta + \phi_e)$ at polar angle θ , with E as the number of error functions, determining the frequency of the border noise, and $\{a_e\}$ and $\{\phi_e\}$ as random amplitudes and offsets. The number

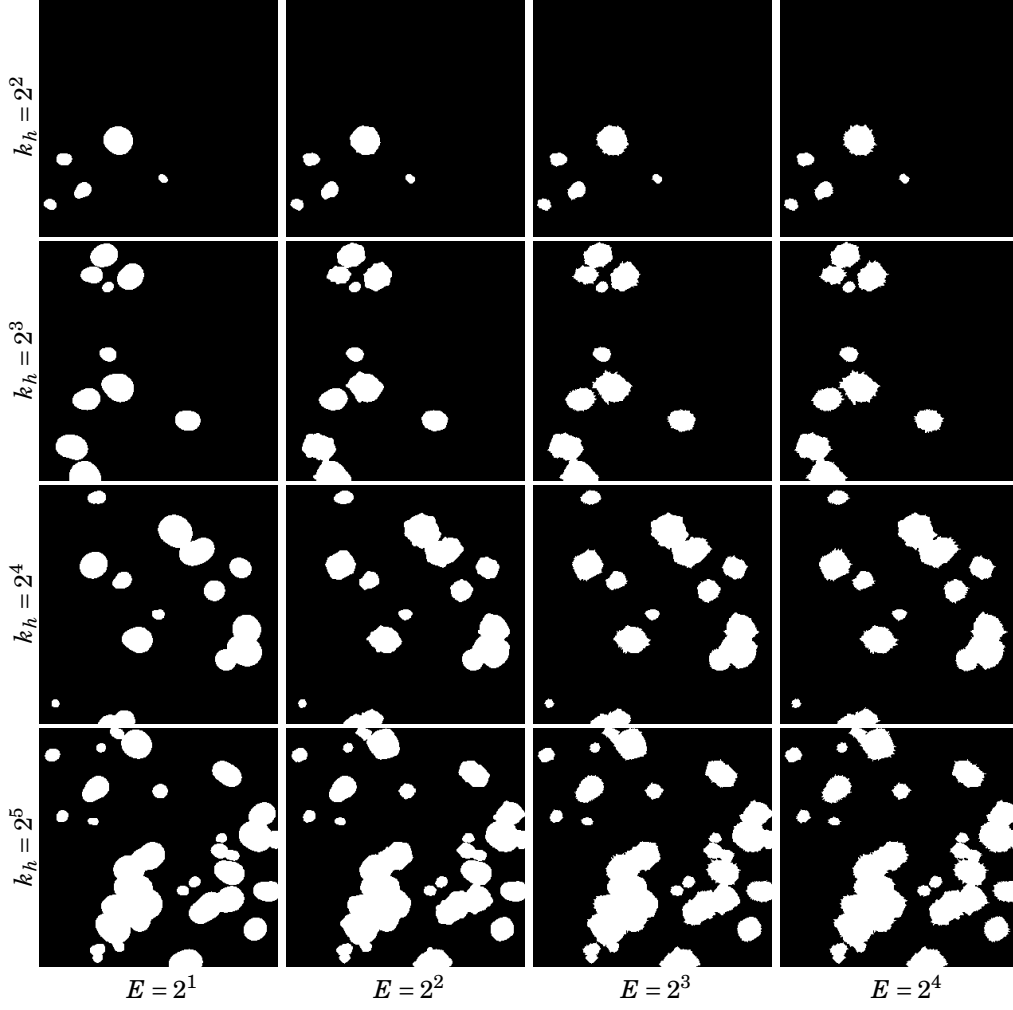


FIGURE 8.6. 300×300 pixel erosion masks, visualizing the impact of the expected rate of hole occurrence k_h and number of error functions E on the synthetic abrasion process.

and irregularity of holes are governed by $n_h \sim \Gamma(k_h)$ and E , respectively. The influence of those parameters can be seen in FIG. 8.6. $k_h = 5$ and $E = 15$ have been selected for our experiments.

By flood-filling the generated holes, which can overlap, of course, we obtain a mask $\mathcal{M}$ which we use to delete surface texture selectively. Instead of leaving the holes as colorless voids, we fill them with a salient green hue which makes them easy to recognize and also allows us to investigate the influence of the texture in these areas on the reconstruction. Applying the mask as it is to the surface would result in unnaturally sharp edges at the borders of the holes. Instead, we use an alpha blending in the vicinity of those edges to smoothen the transitions. To this end, we employ an alpha mask $\mathcal{M}_\alpha = 1 - \exp(-\mathcal{M}_{dist}^2 / 2\sigma_\alpha^2)$, with $\sigma_\alpha = 8$ governing the smoothness. $\mathcal{M}_{dist}$ is the mask $\mathcal{M}$ after a distance transform [307], which gives at each pixel location the distance to the closest zero pixel. The surface image with the synthetic abrasion $\mathcal{I}_{abr} = [\mathcal{I} \cdot (1 - \mathcal{M}_\alpha) + \mathcal{M} \cdot \mathcal{M}_\alpha]$ is obtained by combining the original surface $\mathcal{I}$ with the generated mask $\mathcal{M}$.

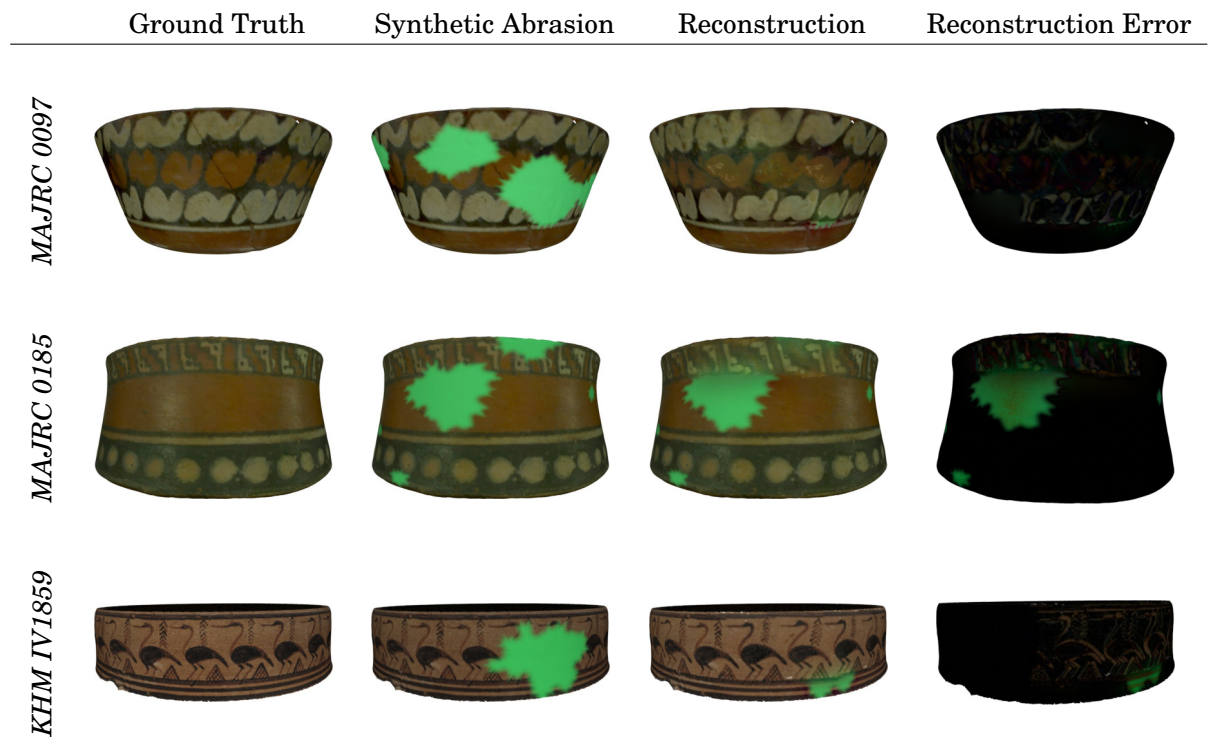


FIGURE 8.7. Three exemplaric pottery objects with their original 3D model (first column), their model after applying the synthetic surface abrasion (second column), after the pattern completion with our method (third column) and the color-coded difference between original and reconstruction (fourth column). The artificial chip-offs are filled with uniform green color.

8.3.3.3 Qualitative Synthetic Experiments

To qualitatively assess the pattern completion after applying a synthetic abrasion, we select three distinct objects with diverse characteristics and thus posing different challenges. They are displayed in FIG. 8.7, left column. The first one is the pre-Columbian bowl shape *MAJRC 0097*, which exhibits three rows of ‘N’-shaped patterns with white and reddish hues. The second object is the pot *MAJRC 0185* with two different pattern types – a very simple circular pattern as well as a triangular staircase pattern, appearing with alternating orientations. The third object, the Attic Geometric high-rimmed bowl *KHM IV1859*² is the most complex input. Although its surface painting is monochromatic (black), it features three distinct pattern classes which appear in a repetitive manner: an organic bird shape, a cross-hatched triangle, and a stack of inverted ‘V’s. All of them appear a total of 26 times around the solid of revolution. The second column in FIG. 8.7 illustrates the inputs after they are subjected to the synthetic surface abrasion. Note that the generated artificial chippings (in green color) vary greatly in shape and size. The third column shows the models from the previous column after applying the automatic pattern completion, while the fourth column constitutes a quality map revealing the differences between the ground

²<https://odeeg.acdh.oeaw.ac.at/vases/object/detail/100>

truth and the reconstruction. These maps are obtained by computing the pixel-wise L_2 norm of the respective RGB color channels.

For *MAJRC 0097* it can be seen that the reconstructed patterns in the top and bottom rows are visually indistinguishable from the existing ones. A slight inaccuracy can be observed with the red patterns in the middle row, but they also appear plausible at large. Note that some of the green color, marking the surface abrasion, is still visible below the bottom row. This can be attributed to the fact that this region is outside the bounding boxes of the patterns and is thus not part of the reconstruction process. Thin outlines of the depicted patterns are perceptible in the quality map, indicating that reconstructed patterns' shapes marginally differ from the originals'. *MAJRC 0185* in the second row features a gap with seven missing patterns of alternating orientation. The completion is able to correctly predict their orientations. However, some ghosting artifacts can be observed. For *KHM IV1859*, four of the bird patterns are reconstructed but no errors could be spotted, neither from the visual inspection nor the quality map.

8.3.3.4 Quantitative Synthetic Experiments

Besides the qualitative assessment in SEC. 8.3.3.3, which shows just one specific surface abrasion use case, we also conduct a quantitative evaluation based on an objective metric. We leverage the quality maps (also shown in FIG. 8.7, right) to compute a single quality measure per experiment. To this end, we integrate over this map and divide the result by the number of pixels that are to be reconstructed and the maximum value of color intensities. This gives a value between 0 and 1, where 0 describes a perfect match and 1 indicates complementary colorization. The same measure is also determined for the synthetically damaged texture prior to the completion. We compute these pairs of measures for 100 different synthetic abrasion experiments for each of the three objects in FIG. 8.7. The resulting measurements plotted over the relative removed surface are shown in FIG. 8.8. The error of the reconstruction is mostly within the range of 0.05 and 0.15 and does not seem to change with an increasing amount of reconstructed surface area. The error between the ground truth and the damaged surface, however, describes a fast ascend before seemingly converging towards about 0.3. This is to be expected, as the green hue is not the ultimate complementary to the original texture.

8.3.4 Discussion, Limitations and Future Work

Apparently, at least some of the green color, marking the synthetic abrasions, is still visible after the completion process (FIG. 8.7). This is due to the fact that our reconstruction is limited to the areas populated by patterns. For most practical examples, this is sufficient (FIG. 8.5) as the clay at the chipped or worn-off regions generally has a similar color to the rest of the vessel. Yet, in some cases, other surface decorations, *e.g.*, straight lines around the solid of revolution, which are without the scope of the proposed method, are also incomplete. For such cases, a processing step

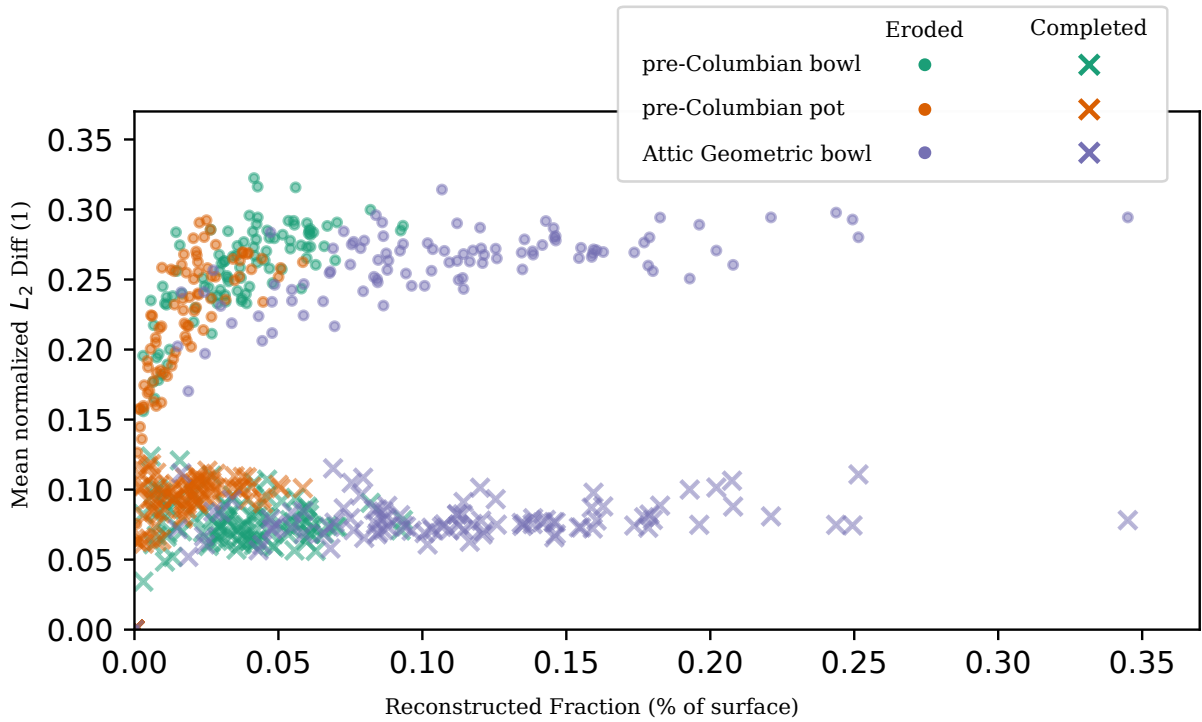


FIGURE 8.8. Reconstruction error (‘x’) and the difference between synthetically damaged surface and ground truth (‘o’), over the relative removed surface area for three different input objects.

combining a chipping detection [228] with context-based inpainting [308], prior to the pattern completion would be a viable addition to our workflow.



FIGURE 8.9. Ghosting artifacts can be observed if a reconstructed pattern and parts of the original painting are not perfectly aligned.

In some cases, ghosting artifacts can be observed occasionally (FIG. 8.9) at the boundaries between original and reconstructed patterns. This is especially the case if the original sequence has

varying distances in-between patterns since we assume a uniform distribution. This assumption of regularity is sometimes problematic, especially in the vicinity of an ornament band's seam, because here the ancient painter had to adjust the patterns' widths to the still available space, making them either narrower or broader. There are three ways to address this issue. First, the stitching algorithm [306] used to fuse the reconstructed patterns with the rest of the surface is not optimal for our workflow as it is intended for the stitching of panoramic photographs. For our application, the inserted texture patches are oftentimes too small, and the overlap with the rest of the surface too insignificant to generate reasonable results. However, – same as with the other parts of our pipeline – this building block can be readily replaced with an algorithm more suited for this task. Other viable remedies include a fine alignment of the patch before it is stitched or the incorporation of the user into the completion process.

Another important question is how many preserved patterns are necessary to reasonably reconstruct a sequence. Generally, the number of preserved patterns needs to be significantly higher than the number of missing ones. Our experiments indicate that completion is feasible with up to 30% of a sequence missing. With even higher fractions, the generation rule detection (SEC. 8.3.1) and gap detection (SEC. 8.3.2.1) are no longer reliable.

8.3.5 Applications

The proposed approach contributes to two significant aspects of archaeological pottery research. First, for the attribution of pottery objects (so-called vases) to painters/workshops and to manufacturing techniques, and second, for educational and communicational purposes regarding these CH objects. Within the field of research on ancient vase painting, the determination of a pattern sequence's generation grammar can be used for recognizing characteristic properties of painters' hands or pottery workshops. For instance, friezes with rows of birds are very common in Geometric pottery. From the automatically determined generation grammar, it is possible to analyze their similarity to other occurrences on different pottery objects and to cluster them accordingly. From a ceramologist's view, the estimation of the number of missing patterns is generally useful as it helps to grasp the painter's overall plan. Another important aspect of this reconstruction approach is the prediction of expected patterns, especially in cases where such repetitive patterns are only partly preserved. As a proxy, similar to the domain knowledge applied by an archaeologist processing pottery, this can strongly support computer-assisted pattern recognition or retrieval tasks. Regarding the educational field, the reconstruction of missing parts is relevant for presenting CH objects – physically as well as virtually, *e.g.*, for teaching activities or for replicas.

8.4 Conclusion

We present an automatic approach for the completion of pattern sequences on ancient pottery. To this end, we determine the sequence’s grammar by inferring a prototypic transformation between consecutive patterns. This generation grammar allows us to detect gaps in the sequence, which will fill with a statistical denoising approach using the sequence’s preserved patterns. The employment of this workflow for real-world examples, as well as synthetically generated use cases, proves the applicability of our approach to the problem at hand. Hence, the contribution of this chapter is the proposed workflow with some self-created and some preexisting building blocks. We want to stress that those could be exchanged and adopted independently to improve certain aspects or address additional challenges. We also note that the problem we solve is very specific and needs to be further extended and combined with other reconstruction techniques (geometric completion, inpainting, *etc.*) for most real-world use cases. However, this is just a first cautious dive into the field of digital restoration, which does not belong to the major topics of the thesis. Still, the presented approach can be a valuable starting point for further research in this direction.

CONCLUSION AND OUTLOOK

Contents

9.1	Reflection on Research Questions	179
9.2	Summary of Contributions	181
9.3	Future Perspective	182
9.4	Final Thoughts	184

In this thesis, I investigated novel approaches for different tasks related to digital archaeology. A particular focus was on interactive retrieval and visualization. This final chapter concludes the thesis with a reflection on the research questions (SEC. 9.1), stated at the beginning; a summarization of the presented contributions (SEC. 9.2); a future outlook (SEC. 9.3), discussing how follow-up-research can build upon the presented techniques; and finally a few closing remarks (SEC. 9.4), which conclude this thesis.

9.1 Reflection on Research Questions

This section explores how the concepts and findings, outlined in the technical chapters (CHAPTER 4–8) of this thesis contribute to answering the guiding research questions formerly stated in SEC. 1.2.1.

RQ1: *“How can user expertise be incorporated interactively in a query formulation for the specification of missing shape information?”*

Regarding **RQ1** it can be noted that the novel query modality of an incomplete 3D shape and a user sketch (3D+Sketch), introduced in SEC. 4.4, provides an intuitive and easy-to-use way

to incorporate a user's shape hypothesis in a query formulation. Both qualitative experiments with domain experts completing a shape, and quantitative experiments with artificially distorted sketches showed, that the proposed method is robust *w.r.t.* noisy sketches. *I.e.*, even a very rough sketch is usually able to improve a retrieval substantially if the overall proportions are on target. While shape and texture-based descriptors proved to be a good choice for the retrieval of CH artifacts, it should be noted that the choice of descriptors is independent from the overall concept and can thus be readily generalized for other domains dealing with P3DOR.

RQ2: *“How can a retrieval approach effectively support the search for specific motifs and patterns, exhibited by the vase paintings?”*

In view of the disambiguation introduced in SEC. 5.1, we split the problem into the retrieval based on motif information and the retrieval based on pattern information. For motifs, a concept has been proposed which combines a semantic segmentation of images with a shape and color-based similarity computation (SEC. 5.3). It has been demonstrated that this method is able to successfully retrieve semantically similar motifs from a motif database built from archaeological photographs. However, the recognition of the relation between motifs or the comprehension of depicted scenes, is a substantially more complex problem and remains an open research objective.

The retrieval of patterns, on the other hand, proved to be more difficult than motifs. This can be attributed to the fact that an expert's notion of similarity for patterns cannot be as easily defined in a formal manner as with motifs, which are generally well-described by their silhouette and color scheme. Nonetheless, progress has been made in this respect. First, with an interactive system for the semi-automatic generation of archaeological drawings (SEC. 6.5). To this end, a self-similarity detection, operating in the background, finds (all) occurrences of a selected pattern and proposes a hypothesis for the drawing of these. A user can confirm or disregard these findings on a per-pattern basis and export the results. Since it was noticed that not much further progress on this problem is possible without the usage of sophisticated deep learning techniques, additional work focused on the generation of annotations and training data, which is a prerequisite for the application of such. That is, we presented two orthogonal approaches to this end. An annotation pipeline and evaluation metric for any type of pattern which appears in a repetitive manner (SEC. 6.3). In a second work we looked into a narrower, more specific problem – the classification of patterns on Geometric pottery (SEC. 6.4) – which was small enough that CNNs were able to provide meaningful results even with a very limited number of training examples.

RQ3: *“How can interactive visualizations support an expert user in discovering (new) inter-object cross-correlations and links?”*

It was established that some interesting cross-correlation are only revealed if different attribute of the objects of a collection are considered simultaneously. *E.g.*, several similar vessel shapes

occurring at the *same* epoch at the *same* location may be an indication of an ancient workshop, while the same observation cannot be made if all vessels of a certain epoch are considered. This kind of exploration is supported by the LVVES system, proposed in SEC. 7.3. Even though no, yet unknown, relations have been discovered, significant relations, known to be present in the dataset used for experiments, have been unveiled automatically through a newly developed correlation measure. Experiments suggested that the proposed system is easily extendable and could also be generalized to other domains.

RQ4: *“How can the generation grammar, intrinsic to repetitive ornaments in vase paintings, be leveraged for the detection and the reconstruction of missing sequence parts?”*

A naïve approach for determining a generation grammar paired with a probabilistic denoising method can be leveraged for the textual completion of incomplete pattern sequences (SEC. 8.3). The narrowness of the problem definition addressed by the proposed concept, as well as some non-negligible limitations, such as the aliasing artifacts in some cases, point out the need for more sophisticated methods for the robust application on real-world objects. Nonetheless, the proposed method showed that viable results are possible even with comparably simple approaches, which can serve as an algorithmic basis.

To sum up, it can be noted that all the stated research questions have been addressed with the technical contributions. **RQ1– RQ3** received undeniably more attention than **RQ4**, which was only opened up later in the project due to a shift in research focus. **RQ1** and **RQ3** (the core objectives of the original project plan) have been dealt with thoroughly, while the conducted experiments show that the contributions for **RQ2** and **RQ4** constitute valuable preparatory work that merits further research effort.

9.2 Summary of Contributions

The main contributions of this thesis are aligned with the above stated research questions and have been published in several peer-reviewed publications. In detail, they include:

- RQ1** • A new query modality (3D+Sketch) which has been subjected to both qualitative and quantitative evaluations [72, 73];
- RQ2** • A new retrieval approach for the figurative information on vase paintings using a combination of semantic segmentation and color and shape-based similarity [80];
 - A benchmark dataset for surface patterns, comprising 82 annotated 3D models with a total of 102 similarity classes and 2,529 annotated patterns [82];
 - An interactive annotation concept for the annotation of Geometric patterns [83];

- A semi-automatic workflow for generating archaeological drawings of repetitive surface patterns, based on an interactive self-similarity detection [81];
- RQ3** • A novel multi-view exploration system (LVVES), allowing the analysis of large data collections along different attributes using parallel links dynamic levels of visual granularity [56]; and
- RQ4** • A pattern completion pipeline for detecting gaps in a pattern sequence and proposing a completion hypothesis [58].

Additionally, several minor algorithmic contributions have been made in the course of working on those topics. The most significant of them are:

- A new unwrapping method for solids of revolution resulting in less distortion in areas of high curvature along the rotation axis through a parametrization of the profile curve (SEC. 3.1);
- A new 3D feature descriptor based on parametrized profile curves (SEC. 4.3);
- A novel graph sparsification algorithm for *fully connected weighted* graphs (ALGORITHM 3), extending the approach by Satuluri *et al.* [277]; and
- A graph correlation measure for weighted graphs with a common set of nodes but different weights (SEC. 7.3.3.3).

9.3 Future Perspective

Future work items, specific to certain proposed systems, are remarked upon in the respective technical chapters. In general, I believe one research direction with great potential is the application of deep learning based techniques for different retrieval tasks. The reason such techniques have not been used more extensively within the scope of this thesis lies in the sparsity of labeled training data. However, towards the close of the project, a large amount of effort was put into the generation of such data, which should provide a good basis for future research efforts in this direction. In hindsight, this should have been done preferably at the start of the project, so that the developed retrieval methods could have been directly evaluated on those benchmarks. Regarding deep learning, I also found that it pays off to downsize the problem to be tackled. *E.g.*, it was observed that CNNs can provide somewhat viable performance for the problem of recognizing Geometric patterns with a rather small set of similarity classes and very homogenous inputs (SEC. 6.4), while preliminarily experiments for the retrieval of arbitrary vessel shapes (SEC. 4.4) pretrained networks were not able to deliver any meaningful results at all. Hence,

I think future concepts could focus on narrower problems, which could be combined hierarchically, *e.g.*, a classification of vessel type (Hellenic, Roman, pre-Columbian, *etc.*), painting style (black-figured, red-figured), and so forth, in order to solve larger more complex problems.

Thematically, I believe the biggest potential for further research lies in the content-based retrieval of vase paintings. While shape-based pottery retrieval has been extensively researched for decades (SEC. 4.2), the retrieval based on vase paintings has been almost left untouched [55], which could possibly be attributed to its much higher complexity. As opposed to vessels shapes, which do not have that many degrees of freedom due to their functional design and method of production, vase paintings were much more varied across cultures and epochs (monochromatic vs. polychromatic paintings or organic vs. geometric shapes, just to name a few). Hence, I think research efforts in this direction should not be overly ambitious but focus on very specific but manageable subproblems, *e.g.*, recognition of just a handful of different patterns/motifs on a specific vase type [221]. *I.e.*, one interesting research question could be whether and how vessel shapes correlate with vase paintings. Since preliminary experiments [230] indeed suggest a connection, this topic certainly merits further research effort.

Resources should also be invested to further develop methods such that they are ready for the market, since the scientific research usually does not go beyond the development of a first prototype for experimental evaluations. On the one hand, that means that, in most cases, the robustness has to be increased significantly, so that arbitrary inputs can be processed successfully. On the other hand, the usage via intuitive interfaces and without the setting of hyperparameters needs to be supported, as it has to be taken into account that CH practitioners are oftentimes not accustomed to computerized workflows and need to be familiarized with scientific software tools. Evans and Daly [309, p.45] remark on this subject:

Archaeologists, it seems, cannot be trusted with equipment that use batteries without breaking something – electronic casualty rates in the field are very high. Contracting units have now become very wary of investiture in technology for the field given its short operating life, and it is unlikely that contracting units will invest in technology now, unless it can be shown clearly that the investment would reap rewards in saving both time and money, and this is generally a good thing.

It is thus unsurprising that Karl *et al.* [55] found in their recent survey that *none* of the presented techniques made its way into the everyday application by domain researchers. During my research, I was also faced with the problem that reference methods were oftentimes not available for comparison as they were only published conceptually or provided by means of an experimental implementation at best. I believe that bringing techniques into a marked-ready state with long-term support would not only greatly contribute to the visibility but also result in them finding their way into the actual use by the intended community.

9.4 Final Thoughts

In summary, all the defined research questions have been addressed. Because of the broadness of topics covered within this project, some of them (motif search, restoration) have been covered in less depth than others (P3DOR, visualization). The already ambitious, initially intended, range of research foci even broadened in the course of the project, due to an opportunistic topic selection, driven by availability of necessary data and incidentally generated ideas, resulting from various experiments. In the hypothetic case that I would have the opportunity to redo the project from scratch, I would take a more narrow approach, pursuing less but therefore well-defined problem cases. On the other hand, the heterogeneity of topics resulted in several ideas for future research and allowed me to immerse myself in different fields of visual computing, such as geometry processing or virtual reality just to name a few, which I would not have experienced otherwise. I believe that the presented techniques and results constitute a significant contribution to the field of digital pottery analysis and sincerely hope that some of them will find their way into the everyday work of archaeologists.



APPENDIX A: MATHEMATICAL NOTATION

As the mathematical descriptions used in this thesis are compiled from many publications which follow different principles, defining a universally applicable notation is a challenging task. Nonetheless, apart from a few minor deviations in some specific cases, the notation is based on the following principles. Note, however, that for the sake of simplicity, the scope of symbols is restricted to the chapter they appear in – *i.e.*, they are oftentimes reused.

X, Y, Z	The coordinate axes X , Y , and Z .
a, α, A	Scalars. [†]
$\mathbf{a}$	Vectors.
α, β, γ	Angles.
$A, \mathcal{A}, \mathbf{A}$	Sets and Matrices.
$A = \{a_0, \dots, a_n\}$	The set A with elements $a_0, \dots, a_n$. [‡]
$\emptyset$	The empty set.
$\mathbb{R}, \mathbb{R}^n, \mathbb{R}^{n \times m}$	The sets of real numbers, n -vectors, and real $n \times m$ -matrices, respectively.
$\mathbb{B}, \mathbb{N}, \mathbb{R}$	The sets of binary numbers $\{0, 1\}$, natural numbers, and real numbers respectively.
$\mathbf{a} = (a_0, \dots, a_n)$	The vector $\mathbf{a}$ with elements $a_0, \dots, a_n$. [‡]
$\mathbf{0}$	The zero vector.
$\langle a_0, a_1 \rangle$	The tuple with elements a_0, a_1 .
$[a_0, a_1], (a_0, a_1], [a_0, a_1), (a_0, a_1)$	The real intervals $\{x \mid a_0 \leq x \leq a_1\}$, $\{x \mid a_0 < x \leq a_1\}$, $\{x \mid a_0 \leq x < a_1\}$, and $\{x \mid a_0 < x < a_1\}$, respectively.
$[a_0 \dots a_1], (a_0 \dots a_1], [a_0 \dots a_1), (a_0 \dots a_1)$	The integer intervals $\{x \mid a_0 \leq x \leq a_1\}$, $\{x \mid a_0 < x \leq a_1\}$, $\{x \mid a_0 \leq x < a_1\}$, and $\{x \mid a_0 < x < a_1\}$, respectively.
$ \mathbf{a} $	The Manhattan norm.
$\ \mathbf{a}\ $	The Euclidean norm.

[†] Scalars are indicated by lowercase Latin and Greek letters and (in rare cases) uppercase Latin letters.

[‡] Indices follow a zero-based numbering.

APPENDIX B: ALGORITHMIC BACKGROUND

Non-Maximum Suppression. A Non-Maximum Suppression (NMS) takes a set of initial detection boxes $\mathcal{B}$ as well as their corresponding confidence scores $\mathcal{S}$ as input and returns the most relevant detection boxes, which are overlapping no more than a given threshold t_{NMS} . The latter is checked with an Intersection over Union (IoU) criterion. The NMS algorithm, in its simplest form [310], is outlined in ALGORITHM 1. The idea is that the given detection boxes are sorted descendingly according to their similarity scores. The first detection box (with the highest similarity) is selected (Line 4) and all remaining boxes exhibiting a too high overlap are removed (Line 7–14). In this manner, we iterate over the whole set.

Algorithm 1 Outline of the Non-Maximum Suppression algorithm.

```

1: function NMS( $\mathcal{B}, \mathcal{S}, t_{NMS}$ )
2:    $\mathcal{D} \leftarrow \{\}$  ▷ The set of relevant candidates
3:   while  $\mathcal{B} \neq \emptyset$  do
4:      $m \leftarrow \mathcal{B}[\text{argmax } \mathcal{S}]$ 
5:      $\mathcal{D} \leftarrow \mathcal{D} \cup \{m\}$ 
6:      $\mathcal{B} \leftarrow \mathcal{B} \setminus \{m\}$ 
7:     for all  $i \in [0..|\mathcal{B}|)$  do ▷ Check for overlapping detection boxes
8:        $b \leftarrow \mathcal{B}[i]$ 
9:        $s \leftarrow \mathcal{S}[i]$ 
10:      if  $iou(m, b) \geq t_{NMS}$  then ▷  $iou$  returns the Intersection over Union
11:         $\mathcal{B} \leftarrow \mathcal{B} \setminus \{b\}$ 
12:         $\mathcal{S} \leftarrow \mathcal{S} \setminus \{s\}$ 
13:      end if
14:    end for
15:  end while
16:  return  $\mathcal{D}$ 
17: end function

```

Graph Aggregation. ALGORITHM 2 outlines the procedure used to aggregate vertices in an undirected weighted graph $G = (V, E)$. The underlying idea is to conduct a greedy pair-wise aggregation of vertices until a desired vertex count $|V| = t_{agg}$ is reached (Line 4). In each iteration, the edge with the lowest global weight is removed from the set of edges (Line 5). Its vertices i, j are removed together with all of their incident edges E_i, E_j (Line 8-10), before a substitute for the removed vertices is added to the set of vertices (Line 12). This new vertex is connected to the other vertices, truthful to the original topology. To this end, we differentiate between vertices which were connected mutually only to either i or j and such which were connected to both of them. In the first case, the edge weights prior to the aggregation can be reused (Line 15-17). In the second case we linearly interpolate between the weights of the edges connecting a vertex to i and j respectively (Line 18-21).

Algorithm 2 Outline of the local node aggregation in undirected weighted graphs.

```

1: function LOCALAGGREGATION( $V, E, t_{agg}$ )
2:    $V_{agg} \leftarrow V; E_{agg} \leftarrow E$ 
3:   Sort  $E_{agg}$  by  $w_{ij}$  ▷ Edges sorted by descending edge weight
4:   while  $|V_{agg}| > t_{agg}$  do
5:      $\langle i, j, w_{ij} \rangle \leftarrow pop(E_{agg})$ 
6:      $E_i \leftarrow \{ \langle i', j', w_{i'j'} \rangle \mid \langle i', j', w_{i'j'} \rangle \in E_{agg} \wedge i' = i \}$ 
7:      $E_j \leftarrow \{ \langle i', j', w_{i'j'} \rangle \mid \langle i', j', w_{i'j'} \rangle \in E_{agg} \wedge j' = j \}$ 
8:      $E_{agg} \leftarrow E_{agg} \setminus \{ \langle i', j', w_{i'j'} \rangle \mid \langle i', j', w_{i'j'} \rangle \in E_i \cup E_j \}$  ▷ Edges incident to  $i$  and  $j$ 
9:     for  $e' \in \{ \langle i', j', w_{i'j'} \rangle \in E_{agg} \mid i' = i \vee i' = j \vee j' = i \vee j' = j \}$  do
10:       $E_{agg} \leftarrow E_{agg} \cup \{e'\}$ 
11:     end for
12:     Let  $v_{agg}$  be the aggregated node for  $i$  and  $j$ 
13:      $V_{agg} \leftarrow V_{agg} \setminus \{i, j\} \cup \{v_{agg}\}$ 
14:      $E'_i \leftarrow \{ \langle i', j', w_{i'j'} \rangle \in E_i \mid (\nexists \langle i'', j'', w_{i''j''} \rangle \in E_j)[j' = j''] \}$ 
15:      $E'_j \leftarrow \{ \langle i', j', w_{i'j'} \rangle \in E_j \mid (\nexists \langle i'', j'', w_{i''j''} \rangle \in E_i)[j' = j''] \}$ 
16:     for  $\langle i', j', w_{i'j'} \rangle \in E'_i \cup E'_j$  do ▷ Mutually exclusive edges
17:        $E_{agg} \leftarrow E_{agg} \cup \{ \langle v_{agg}, j', w_{i'j'} \rangle \}$ 
18:     end for
19:      $V_{common} \leftarrow \{ i' \mid \langle i', j', w_{i'j'} \rangle \in E_i \setminus E'_i \}$  ▷ Vertices connected to  $i$  and  $j$ 
20:     for  $v \in V_{common}$  do
21:        $E_{agg} \leftarrow E_{agg} \cup \{ \langle v_{agg}, v, \frac{1}{2}(w_{iv} + w_{jv}) \rangle \}$ 
22:     end for
23:   end while
24:   return ( $V_{agg}, E_{agg}$ )
25: end function

```

Graph Sparsification. ALGORITHM 3 outlines the procedure used for the local edge sparsification of a graph $G = (V, E)$. The concept is based on the concept proposed by Satuluri *et al.* [277],

Algorithm 2, but adapted for fully-connected undirected weighted graphs.

Algorithm 3 Local Sparsification Algorithm, based on Satuluri *et al.* [277], Algorithm 2, adapted for fully connected weighted graphs.

```

1: function LOCALSPARSIFICATION( $V, E$ )
2:    $E_{sparse} \leftarrow \emptyset$ ;  $d_{min} \leftarrow \infty$ ;  $d_{max} \leftarrow 0$ 
3:   for  $v \in V$  do
4:      $E_v \leftarrow \{\langle i, j, w_{ij} \rangle \in E \mid i = v \vee j = v\}$  ▷ Edges incident to  $v$ 
5:     for  $e = \langle i, j, w_{ij} \rangle \in E_v$  do
6:        $e.sim \leftarrow Sim_V(i, v)$  ▷ Similarity of edge vertices according to EQ. (7.2)
7:     end for
8:     Let  $E_{sim}$  be the edges  $E_v$  sorted by  $e.sim$ 
9:     Let  $E_w$  be the edges  $E_v$  sorted by their weights
10:     $v.d \leftarrow Sim_E(E_{sim}, E_w)$  ▷ The virtual degree of  $v$  according to EQ. (7.3)
11:    if  $v.d < d_{min}$  then  $d_{min} \leftarrow v.d$ 
12:    if  $v.d > d_{max}$  then  $d_{max} \leftarrow v.d$ 
13:  end for
14:  for  $v \in V$  do
15:     $d_{norm} \leftarrow (v.d - d_{min}) / (d_{max} - d_{min})$  ▷ The normalized virtual degree for  $v$ 
16:    Let  $E'_v$  be the top  $f(d_{norm})$  edges of  $E_w$  ▷  $f(d_v)$  according to EQ. (B.1),  $E_w$  as in Line 9
17:     $E_{sparse} \leftarrow E_{sparse} \cup E'_v$ 
18:  end for
19:  return  $(V, E_{sparse})$ 
20: end function

```

The idea is to model the similarity between two vertices connected by an edge by means of the vertices' local neighborhood (Line 6). Satuluri *et al.* [277] base this similarity measure on the simple heuristic that “An edge (i, j) is likely to (not) lie within a cluster if the vertices i and j have adjacency lists with high (low) overlap.” For a fully connected graph, adjacency lists are per-definition equal for all the vertices and are thus no vessel for determining this similarity. Conversely, we base our notion of similarity on the vertices' agreement regarding the distances to all other vertices. *I.e.*, the similarity is high if both vertices have similar distances to all other vertices. Next, in Line 10, we determine a proposed target degree for a vertex. We employ the simple heuristic that if the sequence of edges sorted by similarity is similar to the sequence of edges sorted by their weights, this implies a high vertex degree. The highest and lowest virtual degree over all vertices is also memorized (Line 11–12). In a second loop (Line 14–18) those values are then used to determine the vertex-wise normalized degree which is then finally used to obtain the actual degree of a vertex after the sparsification (Line 15), using the mapping function

$$(B.1) \quad f(d) = \left\lceil \alpha \cdot d^\beta + \gamma \right\rceil$$

This concave mapping ensures that vertices with higher degree are sparsified more aggressively than those with lower degree and that each vertex has at least one edge. In our experiments, we used the empirically determined hyperparameters $\alpha = 1.2$, $\beta = 6.0$, and $\gamma = 1.0$.

LIST OF TABLES

TABLE	Page
2.1 Selection of online data sources	19
2.2 Delineation of photographs obtained from different CVA fascicules	20
2.3 Overview of the used 3D models	25
4.1 Taxonomy of 3D Object Retrieval Methods	47
4.2 Evaluation of the performance of NPC and W-NPC on the SHREC 2021 track	55
6.1 Quantitative performance comparison	115
6.2 Evaluation of different ResNet architectures and hyperparameters	123

LIST OF FIGURES

FIGURE	Page
1.1 Selection of exhibits from the Institute of Classics at the University of Graz	2
1.2 Overview of Sources, Data Modalities, Preparations, and Applications	4
2.1 Exemplaric profile curve as depicted in domain publications	22
2.2 Randomly selected sample of photographs extracted from the CVA and the BAPD . .	23
2.3 Plate Segmentation Pipeline	27
2.4 Exemplaric double page from Kunisch pattern taxonomy	28
3.1 Comparison of cylindrical unwrapping methods	31
3.2 Processing steps of object segmentation	33
3.3 Edge cases limiting automatic object segmentation	33
3.4 Histogram spreading for contrast enhancement	35
3.5 Comparison of instance segmentation approaches applied to different inputs	36
3.6 Comparison of different motif segmentation methods	36
4.1 Design concept of the normalized profile curve descriptor	54
4.2 Precision-recall plots for SHREC track challenges	56
4.3 Confusion matrices of W-NPC for the SHREC track challenges	56
4.4 Proposed workflow for 3D+Sketch-based retrieval	59
4.5 Sketch completion error cases	60
4.6 Comparison of texture completion techniques	61
4.7 Importance weighting with selection grid or brush tool	65
4.8 Local feature similarity of HoG and SCD features	66
4.9 Interface of the prototype implementation	67
4.10 Qualitative retrieval results for <i>KF9049</i>	69
4.11 Qualitative retrieval results for <i>KHM IV 1</i>	72
4.12 Fracturing based on Voronoi tessellation	73
4.13 Synthetic sketch-line distortion concept	74
4.14 Quantitative synthetic fracture experiment with <i>KF9049</i>	75
4.15 Quantitative synthetic fracture experiment with randomly selected query objects . . .	76

5.1	Two vessels exhibiting colorimetric and geometric surface decorations	84
5.2	Visualization of the shape context FD	88
5.3	Motif selection interface	90
5.4	Retrieval result for a person with an outstretched arm	91
5.5	Retrieval result for a palmette leaf	91
5.6	Retrieval result for a depiction of Eros	92
5.7	Additional results for motif-driven retrieval	93
5.8	Limitations of motif-driven retrieval	94
6.1	Pattern recognition benchmark examples	105
6.2	Target dataset for ornament annotation	106
6.3	Different impairments characteristic for scanned pottery	107
6.4	Annotation pipeline for repetitive patterns	108
6.5	Camera setup for capturing sample patterns	109
6.6	Interface of the pattern entity annotation tool	111
6.7	Comparison of different retrieval results	113
6.8	Precision-recall curves and limitation example	115
6.9	Visual exploration system for repetitive patterns	116
6.10	Manual annotation and classification of <i>KHM IV 1</i>	117
6.11	Web-based interface for the annotation of Geometric patterns	120
6.12	Sketch-like input abstraction of 3D surfaces	121
6.13	Precision-recall curves and confusion matrices of different ResNet variants	122
6.14	Surface annotation of <i>KHM IV 1</i>	125
6.15	Surface annotation of <i>Athens KER 2146</i>	126
6.16	Self-Organizing Map showing a subset of the Kunisch patterns collection	129
6.17	Comparison of manual drawing of <i>KFUG G26</i> with a naïve edge detection	131
6.18	Manual creation process of archaeological drawings	132
6.19	Workflow for semi-automated creation of archaeological drawings	132
6.20	Different confidence maps of the example of <i>KHM IV3600</i>	134
6.21	Results of the example <i>KHM IV3600</i>	136
6.22	Results for two Peruvian exhibits from MAJRC	137
7.1	Design concept of the LVVES	146
7.2	Shape Preview Surrogate	151
7.3	Balloon views for expanding MOPs at different recursion depths	152
7.4	LVVES with inter-view links determined by cluster cross-correlation	153
7.5	Showcase for using the SSV and TV	155
7.6	Sensory overload resulting from too many parallel links	157
8.1	Examples of eroded surface patterns	162

8.2	Restoration example with patch-inpainting method	163
8.3	Surface pattern completion pipeline	167
8.4	Selection of reference samples of a staircase pattern	169
8.5	<i>MAJRC 0181</i> and <i>KHM IV3600</i> together with their reconstruction	172
8.6	Comparison of hyperparameters for the generation of synthetic erosion masks	173
8.7	Analysis of the reconstruction on three synthetically eroded surfaces	174
8.8	Quantitative evaluation of reconstruction error	176
8.9	Ghosting artifacts at boundary regions	176

LIST OF ACRONYMS

- 3DK** 3D Knowledge and Distributed Intelligence. [49](#)
- 3DOR** 3D Object Retrieval. [43–47](#), [49](#), [52](#), [53](#), [64](#), [180](#), [184](#), [199](#)
- ANND** Average Nearest-Neighbor Distance. [110](#), [112](#), [114](#), [115](#)
- API** Application Programming Interface. [154](#)
- BAPD** Beazley Archive Pottery Database. [2](#), [3](#), [18–21](#), [23](#), [28](#), [85](#), [154](#), [193](#)
- BCE** Before the Common Era. [2](#), [3](#), [18](#), [24](#), [67](#), [71](#), [156](#), [166](#)
- BIC** Bayesian Information Criterion. [168](#)
- BoVW** Bag-of-Visual-Words. [45](#), [47](#), [52](#), [62](#)
- CBIR** Content-based Image Retrieval. [x](#), [14](#), [41](#), [43](#), [44](#), [59](#), [60](#), [62](#), [71](#), [74](#), [78](#)
- CE** Common Era. [154](#), [166](#)
- CEDD** Color and Edge Directivity Descriptor. [62](#)
- CH** Cultural Heritage. [i](#), [1](#), [3–6](#), [8–12](#), [14](#), [19](#), [20](#), [24](#), [30](#), [43](#), [45](#), [47](#), [50–52](#), [57](#), [79](#), [85](#), [101](#), [103](#), [105](#), [142–144](#), [146](#), [148](#), [157](#), [161](#), [163](#), [165–168](#), [177](#), [180](#), [183](#), [199](#)
- CLAROS** Classical Art Research Online Services. [48](#)
- CNN** Convolutional Neural Network. [11](#), [31](#), [44](#), [46](#), [47](#), [50](#), [62](#), [86](#), [100](#), [101](#), [118–120](#), [127](#), [129](#), [130](#), [180](#), [182](#), [199](#)
- CrossSAVE-CH** Crossmodal Search and Visual Exploration of 3D Cultural Heritage Objects. [ix](#), [1](#), [9](#), [10](#), [12](#), [13](#), [41](#), [161](#)
- CVA** Corpus Vasorum Antiquorum. [3](#), [5](#), [18–21](#), [23](#), [26](#), [32](#), [67](#), [89](#), [154](#), [191](#), [193](#)
- DOM** Document Object Model. [28](#)

- EGBIS** Efficient Graph-Based Image Segmentation. [35–37](#), [89](#), [92](#)
- FCN** Fully Convolutional Network. [31](#)
- FD** Feature Descriptor. [44](#), [51](#), [53](#), [58](#), [60](#), [62](#), [63](#), [65](#), [68](#), [78](#), [85](#), [87](#), [88](#), [93](#), [95](#), [194](#)
- FG/BG** Foreground/Background. [29](#), [32](#), [34](#), [35](#), [48](#), [50](#), [89](#)
- FPN** Feature Pyramid Network. [31](#)
- FT** First-tier. [55](#), [122](#), [123](#)
- FV** Feature Vector. [44](#), [48](#), [51](#), [60](#), [63–65](#), [79](#), [87](#), [102](#), [103](#), [119](#), [121–123](#), [129](#), [148](#)
- FWF** Austrian Science Fund. [10](#)
- GIS** Geographic Information System. [9](#)
- GMM** Gaussian Mixture Model. [34](#), [168](#)
- GMV** Geographic Map Viewer. [146–150](#), [153–157](#)
- GPU** Graphics Processing Unit. [101](#)
- GUI** Graphical UI. [66](#), [67](#), [110](#)
- HoG** Histogram of Oriented Gradients. [50](#), [62–66](#), [68–70](#), [77](#), [78](#), [193](#), [199](#)
- IoU** Intersection over Union. [111](#), [135](#), [187](#)
- IR** Information Retrieval. [43](#)
- JSON** JavaScript Object Notation. [107](#), [108](#), [110](#), [114](#)
- KFU** Karl Franzens Universität Graz. [24](#), [25](#)
- KHM** Kunsthistorisches Museum Wien. [24](#), [25](#), [67](#)
- LBP** Local Binary Pattern. [101](#), [199](#)
- LD-SIFT** Local Depth SIFT. [51](#), [52](#)
- LIMC** Lexicon Iconographicum Mythologiae Classicae. [3](#)
- LMK** Landesmuseum Kärnten. [24](#), [25](#), [67](#)
- LVVES** Linked Views Visual Exploration System. [146](#), [148](#), [153](#), [154](#), [156](#), [157](#), [181](#), [182](#), [194](#)

- MAJRC** Museo de Arqueología Josefina Ramos de Cox. [24](#), [25](#), [99](#), [106](#), [162](#), [171](#), [194](#)
- mAP** Mean Average Precision. [55](#), [110–115](#), [122](#), [123](#)
- MI-HoG** Mirror Reflection Invariant HoG. [64](#)
- ML** Machine Learning. [46](#), [62](#), [78](#), [85](#), [100–103](#), [118](#), [138](#)
- MMSE** Mininum Mean Squared Error. [170](#)
- MOP** Multi-Object Preview. [146](#), [147](#), [149–153](#), [194](#)
- mpLBP** Mean Point LBP. [101](#)
- MPEG** Moving Picture Experts Group. [50](#), [51](#)
- nDCG** Normalized Discounted Cumulative Gain. [55](#)
- NMS** Non-Maximum Suppression. [11](#), [73](#), [134](#), [135](#), [187](#)
- NN** Nearest Neighbor. [55](#), [122](#), [123](#)
- NP** Non-deterministic Polynomial-time. [148](#)
- NPC** Normalized Profile Curve. [53–57](#), [79](#), [191](#), [193](#), [200](#)
- NURB** Non-uniform Rational B-spline. [49](#)
- OCR** Optical Character Recognition. [26](#), [28](#)
- P3DOR** Partial 3DOR. [45](#), [46](#), [52](#), [53](#), [180](#), [184](#)
- PCA** Principal Component Analysis. [44](#)
- PRESIOUS** PREdictive digitization, reStoration and degradatIOon assessment of CH objectS. [165](#)
- PRISMA** Pattern Recognition, Indexing and Social Media Analysis. [10](#)
- QBE** Query by Example. [43](#), [84–86](#), [102](#)
- R-CNN** Region-based CNN. [31](#)
- ResNet** Residual Neural Network. [50](#), [120](#), [121](#)
- RIFT** Rotation-Invariant Feature Transform. [60](#)
- RLXOR** Randomized Local XOR. [113–115](#)

RND Random Selection. [113–115](#)

ROD Repetitive Ornament Detection. [113–115](#)

SCD Shape Contour Descriptor. [62](#), [63](#), [65](#), [66](#), [68–70](#), [77](#), [78](#), [148](#), [193](#)

SHREC Shape Retrieval Contest. [13](#), [52](#), [55](#), [101–104](#)

SIFT Scale-Invariant Feature Transform. [49](#), [51](#), [52](#), [62](#), [198](#)

SIMA Systeme d’Imagerie et d’Analyse du Mobilier Archeologique. [48](#)

SOM Self-Organizing Map. [129](#), [194](#)

SOP Single Object Preview. [116](#), [146](#), [147](#)

SPS Shape Preview Surrogate. [150–153](#), [155](#), [194](#)

SSV Shape Similarity Viewer. [146](#), [148–151](#), [153–157](#), [194](#)

ST Second-tier. [55](#), [122](#), [123](#)

TV Timeline Viewer. [146](#), [148](#), [149](#), [153–155](#), [194](#)

UI User Interface. [118](#), [120](#), [136](#), [198](#)

UMJ Universalmuseum Joanneum. [24](#), [25](#)

W-NPC Weighted NPC. [55–57](#), [79](#), [191](#), [193](#)

BIBLIOGRAPHY

- [1] Gaffney, V., Baldwin, E., Bates, M., Bates, C. R., Gaffney, C., Hamilton, D., Kinnaird, T., Neubauer, W., Yorston, R., Allaby, R., “A massive, late neolithic pit structure associated with durrington walls henge”, *Internet Archaeology*, 2020 (cit. on p. 1).
- [2] Hulusic, V., Gusia, L., Luci, N., Smith, M., “Tangible Interfaces for VR Cultural Heritage Application - School House Virtual Museum”, in *Eurographics Workshop on Graphics and Cultural Heritage*, The Eurographics Association, 2021. DOI: [10.2312/gch.20211404](https://doi.org/10.2312/gch.20211404) (cit. on p. 2).
- [3] Shitut, K., Geigel, J., Decker, J., Jacobs, G., Doherty, A., “Triggering the Past: Cultural Heritage Interpretation Using Augmented and Virtual Reality at a Living History Museum”, in *Eurographics Workshop on Graphics and Cultural Heritage*, The Eurographics Association, 2021. DOI: [10.2312/gch.20211406](https://doi.org/10.2312/gch.20211406) (cit. on p. 2).
- [4] Rizvic, S., Boskovic, D., Okanovic, V., Ivkovic-Kihic, I., Sljivo, S., “Virtual reality experience of sarajevo war heritage.”, in *GCH*, 2019, pp. 1–6 (cit. on p. 2).
- [5] Anichini, F., Banterle, F., Buxeda i Garrigós, J., Callieri, M., Dershowitz, N., Dubbini, N., Diaz, D. L., Evans, T., Gattiglia, G., Green, K., “Developing the archaide application: A digital workflow for identifying, organising and sharing archaeological pottery using automated image recognition”, *Internet Archaeology*, 2020, num. 52, 2020 (cit. on pp. 2, 10, 46, 47, 51).
- [6] Papaioannou, G., Schreck, T., Andreadis, A., Mavridis, P., Gregor, R., Sipiran, I., Vardis, K., “From reassembly to object completion - a complete systems pipeline”, *Computing and Cultural Heritage*, vol. 10, no. 2, Feb. 2017 (cit. on pp. 2, 3, 7, 9, 12).
- [7] Shen, Y., Wang, Z., Sun, Q., Chen, A., Rushmeier, H., “Reconstructing Dura-Europos From Sparse Photo Collections Using Deep Contour Extraction”, in *Eurographics Workshop on Graphics and Cultural Heritage*, The Eurographics Association, 2021. DOI: [10.2312/gch.20211408](https://doi.org/10.2312/gch.20211408) (cit. on p. 2).
- [8] (2022). Beazley archive pottery database (bapd), [Online]. Available: <https://www.beazley.ox.ac.uk/carc/pottery> (cit. on pp. 2, 3, 18, 19, 48).
- [9] Boardman, J., *Athenian Black Figure Vases*. Thames & Hudson, 1974 (cit. on p. 2).

- [10] Eschbach, N., Schmidt, S., *Töpfer Maler Werkstatt: Zuschreibungen in der griechischen Vasenmalerei und die Organisation antiker Keramikproduktion*, ser. Beihefte zum Corpus vasorum antiquorum ; Bd. 7. CH Beck, 2016, ISBN: 3406669409 (cit. on pp. 2, 22, 42, 83).
- [11] Boardman, J., *Athenian Red Figure Vases: The Archaic Period*. Thames & Hudson, 1975 (cit. on p. 2).
- [12] Boardman, J., *Athenian Red Figure Vases: The Classical Period*. Thames & Hudson, 1989 (cit. on p. 2).
- [13] Johácz, S., “A new method in attribution? attempts of the employment of geometric morphometrics in the attribution of late archaic attic lekythoi”, 2018 (cit. on pp. 2, 50, 68).
- [14] (2022). Lexicon Iconographicum Mythologiae Classicae, [Online]. Available: <http://www.limc.ch> (cit. on p. 3).
- [15] Cook, R. M., *Greek painted pottery, Third edition*. Routledge, 1997 (cit. on pp. 3, 155).
- [16] Boardman, J., *The history of Greek vases: potters, painters, and pictures*. Thames & Hudson, 2001 (cit. on p. 3).
- [17] Kunisch, N., *Ornamente geometrischer Vasen. Ein Kompendium*. Cologne, Weimar, Vienna: Böhlau, 1998 (cit. on pp. 3, 11, 20, 26–28, 99, 117, 119, 124).
- [18] (2022). Cva online, [Online]. Available: <https://www.cvaonline.org/cva/> (cit. on pp. 3, 5, 18, 19).
- [19] Larue, F., Di Benedetto, M., Dellepiane, M., Scopigno, R., “From the digitization of cultural artifacts to the web publishing of digital 3d collections: An automatic pipeline for knowledge sharing.”, *Journal of multimedia*, vol. 7, no. 2, 2012 (cit. on p. 3).
- [20] Razdan, A., Liu, D., Bae, M., Zhu, M., Simon, A., Farin, G., Henderson, M., “Using geometric modeling for archiving and searching 3d archaeological vessels”, in *In CISST*, 2001 (cit. on pp. 3, 7, 10, 46, 47, 49).
- [21] Pintus, R., Pal, K., Yang, Y., Weyrich, T., Gobbetti, E., Rushmeier, H., “A survey of geometric analysis in cultural heritage”, *Computer Graphics Forum*, vol. 35, no. 1, pp. 4–31, 2016. DOI: [10.1111/cgf.12668](https://doi.org/10.1111/cgf.12668) (cit. on pp. 3, 6–8).
- [22] Rasheed, N. A., Nordin, M. J., “A survey of computer methods in reconstruction of 3d archaeological pottery objects”, *International Journal of Advanced Research*, vol. 3, no. 3, pp. 712–714, 2015 (cit. on pp. 3, 7, 9, 162).
- [23] Andreadis, A., Papaioannou, G., Mavridis, P., “Generalized digital reassembly using geometric registration”, in *2015 Digital Heritage*, IEEE, vol. 2, 2015, pp. 549–556 (cit. on pp. 3, 7, 9, 12, 162, 164).

-
- [24] Sipiran, I., “Analysis of partial axial symmetry on 3d surfaces and its application in the restoration of cultural heritage objects”, in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2017, pp. 2925–2933 (cit. on pp. [3](#), [7](#), [9](#), [12](#), [59](#), [162](#), [165](#)).
- [25] Sipiran, I., “Completion of cultural heritage objects with rotational symmetry”, in *Proceedings of the 11th Eurographics Workshop on 3D Object Retrieval*, 2018, pp. 87–93 (cit. on pp. [3](#), [7](#), [9](#), [12](#), [162](#), [165](#)).
- [26] Sharf, A., Alexa, M., Cohen-Or, D., “Context-based surface completion”, in *ACM SIGGRAPH 2004 Papers*, ser. SIGGRAPH ’04, Los Angeles, California: Association for Computing Machinery, 2004, pp. 878–887. DOI: [10.1145/1186562.1015814](#) (cit. on pp. [3](#), [7](#), [9](#), [12](#), [162](#), [164](#)).
- [27] Harary, G., Tal, A., Grinspun, E., “Context-based coherent surface completion”, *ACM Trans. Graph.*, vol. 33, no. 1, pp. 1–12, Feb. 2014. DOI: [10.1145/2532548](#) (cit. on pp. [3](#), [7](#), [9](#), [12](#), [162](#), [164](#)).
- [28] Harary, G., Tal, A., Grinspun, E., “Feature-preserving surface completion using four points”, in *Computer Graphics Forum*, Wiley Online Library, vol. 33, 2014, pp. 45–54 (cit. on pp. [3](#), [7](#), [9](#), [12](#), [162](#), [164](#)).
- [29] Assael, Y., Sommerschield, T., Shillingford, B., Bordbar, M., Pavlopoulos, J., Chatzipanagiotou, M., Androutsopoulos, I., Prag, J., Freitas, N., “Restoring and attributing ancient texts using deep neural networks”, *Nature*, vol. 603, no. 7900, pp. 280–283, 2022 (cit. on pp. [3](#), [7](#), [9](#), [166](#)).
- [30] Aliaga, D. G., Law, A. J., Yeung, Y. H., “A virtual restoration stage for real-world objects”, in *ACM SIGGRAPH Asia 2008 Papers*, ser. SIGGRAPH Asia ’08, Singapore: Association for Computing Machinery, 2008. DOI: [10.1145/1457515.1409102](#) (cit. on pp. [3](#), [7](#), [9](#), [165](#), [166](#)).
- [31] Gilad-Glickman, D., Shimshoni, I., “Color Restoration of Scanned Archaeological Artifacts with Repetitive Patterns”, in *Proc. of the 14th Eurographics Workshop on Graphics and Cultural Heritage*, ser. GCH ’16, Genova, Italy: Eurographics Association, 2016, pp. 105–108, ISBN: 9783038680116 (cit. on pp. [3](#), [7](#), [9](#), [165](#), [166](#)).
- [32] Sablatnig, R., Menard, C., “3d reconstruction of archaeological pottery using profile primitives”, in *Proc. of Intl. Workshop on Synthetic-Natural Hybrid Coding and Three-Dimensional Imaging*, Citeseer, 1997, pp. 93–96 (cit. on pp. [3](#), [7](#), [46–48](#), [53](#)).
- [33] Kampel, M., Sablatnig, R., “Rule based system for archaeological pottery classification”, *Pattern Recognition Letters*, vol. 28, no. 6, pp. 740–747, 2007 (cit. on pp. [3](#), [7](#), [10](#), [46–48](#), [53](#)).

- [34] Martínez-Carrillo, A. L., Lucena, M. J., Fuertes, J. M., Ruiz, A., “Morphometric analysis applied to the archaeological pottery of the valley of guadaluquivir”, in *Morphometrics for Nonmorphometricians*, Springer, 2010, pp. 307–323 (cit. on pp. [3](#), [7](#), [46–48](#)).
- [35] Maiza, C., Gaildrat, V., “Automatic classification of archaeological potsherds”, in *The 8th International Conference on Computer Graphics and Artificial Intelligence, 3IA*, 2005, pp. 2005–12 (cit. on pp. [3](#), [7](#), [46](#), [47](#), [49](#), [53](#)).
- [36] Schurmans, U. A., Razdan, A., Simon, A., Marzke, M., McCartney, P., Van Alfen, D., Jones, G., Zhu, M., Liu, D., Bae, M., “Advances in geometric modeling and feature extraction on pots, rocks and bones for representation and query via the internet”, 2002 (cit. on pp. [3](#), [7](#), [10](#), [46](#), [47](#), [49](#)).
- [37] Femiani, J. C., Razdan, A., Farin, G., “Curve shapes: Comparison and alignment”, *submitted to TPAMI*, 2004 (cit. on pp. [3](#), [7](#), [10](#), [46](#), [47](#)).
- [38] Hristov, V., Agre, G., “A software system for classification of archaeological artefacts represented by 2d plans”, *Cybern. Inf. Technol.*, vol. 13, pp. 82–96, 2013 (cit. on pp. [3](#), [7](#), [45–47](#), [49](#)).
- [39] Piccoli, C., Aparajeya, P., Papadopoulos, G. T., Bintliff, J., Leymarie, F. F., Bes, P., Enden, M., Poblome, J., Daras, P., “Towards the automatic classification of pottery sherds: Two complementary approaches”, *Across Space and Time*, p. 463, 2015 (cit. on pp. [3](#), [7](#), [46](#), [47](#), [49](#)).
- [40] Kurtz, D., Parker, G., Shotton, D., Klyne, G., Schroff, F., Zisserman, A., Wilks, Y., “Claros - bringing classical art to a global public”, in *2009 Fifth IEEE International Conference on e-Science*, 2009, pp. 20–27. DOI: [10.1109/e-Science.2009.11](#) (cit. on pp. [3](#), [7](#), [18](#), [46–48](#), [53](#)).
- [41] Wilczek, J., Monna, F., Navarro, N., Chateau-Smith, C., “A computer tool to identify best matches for pottery fragments”, *Journal of Archaeological Science: Reports*, vol. 37, p. 102891, 2021 (cit. on pp. [3](#), [7](#), [46](#), [47](#), [49](#)).
- [42] Sipiran, I., Lazo, P., Lopez, C., Jimenez, M., Bagewadi, N., Bustos, B., Dao, H., Gangisetty, S., Hanik, M., Ho-Thi, N.-P., Holenderski, M., Jarnikov, D., Labrada, A., Lengauer, S., Licandro, R., Nguyen, D.-H., Nguyen-Ho, T.-L., Perez Rey, L. A., Pham, B.-D., Pham, M.-K., Preiner, R., Schreck, T., Trinh, Q.-H., Tonnaer, L., von Tycowicz, C., Vu-Le, T.-A., “Shrec 2021: Retrieval of cultural heritage objects”, *Computers & Graphics*, vol. 100, pp. 1–20, 2021. DOI: [10.1016/j.cag.2021.07.010](#) (cit. on pp. [3](#), [7](#), [9](#), [12–14](#), [24](#), [42](#), [46](#), [47](#), [52](#), [53](#), [55](#), [56](#), [79](#), [122](#)).

- [43] Novak, J., Micheel, I., Melenhorst, M., Wieneke, L., Düring, M., Morón, J. G., Pasini, C., Tagliasacchi, M., Fraternali, P., “Histogram—a visualization tool for collaborative analysis of networks from historical social multimedia collections”, in *2014 18th International Conference on Information Visualisation*, IEEE, 2014, pp. 241–250 (cit. on pp. 3, 11, 143).
- [44] Preiner, R., Schmidt, J., Krösl, K., Schreck, T., Mistelbauer, G., “Augmenting Node-Link Diagrams with Topographic Attribute Maps”, *Computer Graphics Forum*, 2020. DOI: [10.1111/cgf.13987](https://doi.org/10.1111/cgf.13987) (cit. on pp. 3, 6, 9, 11, 143).
- [45] Maaten, L., Boon, P., Lange, G., Paijmans, H., Postma, E., “Computer vision and machine learning for archaeology”, *Proceedings of Computer Applications and Quantitative Methods in Archaeology*, pp. 112–130, 2006 (cit. on pp. 3, 6, 11, 143, 148).
- [46] Bogacz, B., Feldmann, F., Prager, C., Mara, H., “Visualizing networks of maya glyphs by clustering subglyphs”, 2018 (cit. on pp. 3, 11, 143).
- [47] Roman-Rangel, E., Pallan, C., Odobez, J.-M., Gatica-Perez, D., “Analyzing ancient maya glyph collections with contextual shape descriptors”, *International Journal of Computer Vision*, vol. 94, no. 1, pp. 101–117, 2011 (cit. on pp. 3, 11, 143).
- [48] Deufemia, V., Paolino, L., Tortora, G., Traverso, A., Mascardi, V., Ancona, M., Martelli, M., Bianchi, N., De Lumley, H., “Investigative analysis across documents and drawings: Visual analytics for archaeologists”, in *Proceedings of the International Working Conference on Advanced Visual Interfaces*, ser. AVI ’12, Capri Island, Italy: Association for Computing Machinery, 2012, pp. 539–546. DOI: [10.1145/2254556.2254658](https://doi.org/10.1145/2254556.2254658) (cit. on pp. 3, 11, 143).
- [49] Coburn, J., “I dont know what im looking for: Better understanding public usage and behaviours with tyne & wear archives & museums online collections”, *MW2016: Museums and the Web*, 2016 (cit. on pp. 3, 11, 143).
- [50] Windhager, F., Salisu, S., Leite, R. A., Filipov, V., Miksch, S., Schreder, G., Mayr, E., “Many views are not enough: Designing for synoptic insights in cultural collections”, *IEEE Computer Graphics and Applications*, vol. 40, no. 3, pp. 58–71, 2020 (cit. on pp. 3, 6, 9, 11, 130, 143).
- [51] Simon, R., Isaksen, L., Barker, E. T. E., Soto Cañamares, P., “Peripleo: A tool for exploring heterogenous data through the dimensions of space and time”, *Code4Lib Journal*, 2016 (cit. on pp. 3, 11, 143).
- [52] Scopigno, R., Callieri, M., Cignoni, P., Corsini, M., Dellepiane, M., Ponchio, F., Ranzuglia, G., “3d models for cultural heritage: Beyond plain visualization”, *Computer*, vol. 44, no. 7, pp. 48–55, 2011 (cit. on pp. 3, 18).
- [53] Bloesch, H., “Conseils pour photographier les vases grecs”, in *Colloque international sur le Corpus vasorum antiquorum*, Lyon, 3–5 juillet 1956, Lyon: Centre national de la recherche scientifique, 1957, pp. 31–33 (cit. on pp. 5, 24, 33, 58, 62).

- [54] Mommsen, H., “Zur Auswertung von Formfotos und Profilzeichnungen”, in *Vasenforschung und Corpus Vasorum Antiquorum: Standortbestimmung und Perspektiven*, ser. Corpus Vasorum Antiquorum, Deutschland, Beiheft 1, Munich: Bavarian Academy of Sciences and Humanities, C.H.Beck, 2002, pp. 23–36 (cit. on pp. [5](#), [24](#), [33](#), [58](#), [62](#)).
- [55] Karl, S., Houska, P., Lengauer, S., Haring, J., Trinkl, E., Preiner, R., “Advances in digital pottery analysis”, *it - Information Technology*, 2022. DOI: [doi:10.1515/itit-2022-0006](#) (cit. on pp. [6](#), [7](#), [9](#), [12](#), [13](#), [22](#), [29](#), [34](#), [42](#), [85](#), [183](#)).
- [56] Lengauer, S., Komar, A., Karl, S., Trinkl, E., Preiner, R., Schreck, T., “Visual Exploration of Cultural Heritage Collections with Linked Spatiotemporal, Shape and Metadata Views”, in *Vision, Modeling, and Visualization*, The Eurographics Association, 2020. DOI: [10.2312/vmv.20201196](#) (cit. on pp. [6](#), [9](#), [12](#), [13](#), [15](#), [130](#), [142](#), [182](#)).
- [57] Barthel, K. U., “Improved image retrieval using automatic image sorting and semi-automatic generation of image semantics”, in *2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services*, IEEE, 2008, pp. 227–230 (cit. on pp. [6](#), [130](#)).
- [58] Lengauer, S., Preiner, R., Sipiran, I., Karl, S., Trinkl, E., Bustos, B., Schreck, T., “Context-based Surface Pattern Completion of Ancient Pottery”, in *Eurographics Workshop on Graphics and Cultural Heritage*, The Eurographics Association, 2022. DOI: [10.2312/gch.20221234](#) (cit. on pp. [7](#), [9](#), [12](#), [13](#), [15](#), [162](#), [182](#)).
- [59] Martínez-Carrillo, A., Ruiz, A., Lucena, M. J., Fuertes, J. M., “Computer tools for archaeological reference collections: The case of the ceramics of the iberian period from andalusia (spain)”, in *International Workshop on Multimedia for Cultural Heritage*, Springer, 2011, pp. 51–62 (cit. on pp. [7](#), [21](#), [46–48](#), [141](#), [144](#)).
- [60] Koutsoudis, A., Pavlidis, G., Liami, V., Tsiafakis, D., Chamzas, C., “3d pottery content-based retrieval based on pose normalisation and segmentation”, *Journal of Cultural Heritage*, vol. 11, no. 3, pp. 329–338, 2010 (cit. on pp. [7](#), [10](#), [46](#), [47](#), [51](#)).
- [61] Hörr, C., Brunner, D., Brunnert, G., “Feature extraction on axially symmetric pottery for hierarchical classification”, *Computer-Aided Design and Applications*, vol. 4, no. 1-4, pp. 375–384, 2007 (cit. on pp. [7](#), [46](#), [47](#), [51](#)).
- [62] Koutsoudis, A., Chamzas, C., “3d pottery shape matching using depth map images”, *Journal of cultural heritage*, vol. 12, no. 2, pp. 128–133, 2011 (cit. on pp. [7](#), [46](#), [47](#), [51](#)).
- [63] Sfikas, K., Pratikakis, I., Koutsoudis, A., Savelonas, M., Theoharis, T., “Partial matching of 3d cultural heritage objects using panoramic views”, *Multimedia Tools and Applications*, vol. 75, no. 7, pp. 3693–3707, 2016 (cit. on pp. [7](#), [46](#), [47](#), [51](#)).

- [64] Roman-Rangel, E., Jimenez-Badillo, D., Aguayo-Ortiz, E., “Categorization of aztec potsherds using 3d local descriptors”, in *Asian Conference on Computer Vision*, Springer, 2014, pp. 567–582 (cit. on pp. 7, 42, 46, 47, 52, 62).
- [65] Roman-Rangel, E., Jimenez-Badillo, D., Marchand-Maillet, S., “Classification and retrieval of archaeological potsherds using histograms of spherical orientations”, *Journal on Computing and Cultural Heritage (JOCCH)*, vol. 9, no. 3, pp. 1–23, 2016 (cit. on pp. 7, 46, 47, 52, 62).
- [66] Jimenez-Badillo, D., Roman-Rangel, E., “Application of the ‘bag of words’ model (bow) for analysing archaeological potsherds”, *CAA2015*, p. 847, (cit. on pp. 7, 46, 47, 52, 62).
- [67] Savelonas, M. A., Pratikakis, I., Sfikas, K., “Fisher encoding of differential fast point feature histograms for partial 3d object retrieval”, *Pattern Recognition*, vol. 55, pp. 114–124, 2016 (cit. on pp. 7, 45–47, 52).
- [68] Pratikakis, I., Savelonas, M., Arnaoutoglou, F., Ioannakis, G., Koutsoudis, A., Theoharis, T., Tran, M., Nguyen, V., Pham, V., Nguyen, H., “Shrec’16 track: Partial shape queries for 3d object retrieval”, *Proc. 3DOR*, vol. 1, no. 8, 2016 (cit. on pp. 7, 47, 52).
- [69] Li, B., Lu, Y., Godil, A., Schreck, T., Bustos, B., Ferreira, A., Furuya, T., Fonseca, M., Johan, H., Matsuda, T., Ohbuchi, R., Pascoal, P., Saavedra, J., “A comparison of methods for sketch-based 3D shape retrieval”, *Computer Vision and Image Understanding*, vol. 119, pp. 57–80, Feb. 2014 (cit. on pp. 7, 45).
- [70] Wang, F., Kang, L., Li, Y., “Sketch-based 3d shape retrieval using convolutional neural networks”, in *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015, Boston, MA, USA, June 7-12, 2015*, 2015, pp. 1875–1883. DOI: [10.1109/CVPR.2015.7298797](https://doi.org/10.1109/CVPR.2015.7298797) (cit. on pp. 7, 45).
- [71] Lee, J., Funkhouser, T. A., “Sketch-based search and composition of 3d models.”, in *SBM*, 2008, pp. 97–104 (cit. on pp. 7, 45).
- [72] Lengauer, S., Komar, A., Labrada, A., Karl, S., Trinkl, E., Preiner, R., Bustos, B., Schreck, T., “Sketch-Aided Retrieval of Incomplete 3D Cultural Heritage Objects”, in *Eurographics Workshop on 3D Object Retrieval*, The Eurographics Association, 2019, pp. 17–24. DOI: [10.2312/3dor.20191057](https://doi.org/10.2312/3dor.20191057) (cit. on pp. 7, 9, 12–14, 42, 46, 47, 181).
- [73] Lengauer, S., Komar, A., Labrada, A., Karl, S., Trinkl, E., Preiner, R., Bustos, B., Schreck, T., “A sketch-aided retrieval approach for incomplete 3d objects”, *Computers & Graphics*, vol. 87, pp. 111–122, 2020. DOI: [10.1016/j.cag.2020.02.001](https://doi.org/10.1016/j.cag.2020.02.001) (cit. on pp. 7, 9, 12–14, 42, 46, 47, 181).

- [74] Karl, S., Bayer, P., Márton, A., Mara, H., “Advanced Documentation Methods in Studying Corinthian Black-figure Vase Painting”, in *Proc. of the 23rd Int. Conf. on Cultural Heritage and New Technologies 2018*, Vienna: Museen der Stadt Wien – Stadtarchäologie, 2019, ISBN: 978-3-200-06576-5 (cit. on pp. [8](#), [19](#), [24](#), [30](#), [84](#), [99](#), [100](#), [130](#), [137](#)).
- [75] Bechtold, S., Krömker, S., Mara, H., Kratzmüller, B., “Rollouts of fine ware pottery using high resolution 3d meshes”, in *Proceedings of the 11th International Symposium on Virtual Reality, Archaeology and Cultural Heritage*, ser. VAST '10, Paris: Eurographics Association, 2010, pp. 79–86 (cit. on pp. [8](#), [30](#)).
- [76] Karras, G., Patias, P., Petsa, E., “Digital monoplotting and photo-unwrapping of developable surfaces in architectural photogrammetry”, *International Archives of photogrammetry and Remote Sensing*, vol. 31, pp. 290–294, 1996 (cit. on pp. [8](#), [30](#), [31](#), [109](#), [168](#)).
- [77] Preiner, R., Karl, S., Bayer, P., Schreck, T., “Elastic Flattening of Painted Pottery Surfaces”, in *Eurographics Workshop on Graphics and Cultural Heritage*, The Eurographics Association, 2018. DOI: [10.2312/gch.20181355](#) (cit. on pp. [8](#), [30](#), [131](#), [168](#)).
- [78] Houska, P., Lengauer, S., Karl, S., Preiner, R., “Direct Elastic Unrollings of Painted Pottery Surfaces from Sparse Image Sets”, in *Eurographics Workshop on Graphics and Cultural Heritage*, The Eurographics Association, 2021. DOI: [10.2312/gch.20211417](#) (cit. on pp. [8](#), [9](#), [12](#), [13](#), [30](#), [94](#)).
- [79] Lee, Y. J., Zitnick, C. L., Cohen, M. F., “ShadowDraw: real-time user guidance for freehand drawing”, *ACM Trans. Graph.*, vol. 30, no. 4, 27:1–27:10, 2011. DOI: [10.1145/1964921.1964922](#) (cit. on pp. [8](#), [78](#), [120](#)).
- [80] Lengauer, S., Komar, A., Labrada, A., Karl, S., Trinkl, E., Preiner, R., Bustos, B., Schreck, T., “Motif-driven retrieval of greek painted pottery”, in *Eurographics Workshop on Graphics and Cultural Heritage*, Eurographics-European Association for Computer Graphics, The Eurographics Association, 2019, pp. 91–98. DOI: [10.2312/gch.20191354](#) (cit. on pp. [8](#), [9](#), [12–14](#), [84](#), [181](#)).
- [81] Lengauer, S., Komar, A., Karl, S., Trinkl, E., Sipiran, I., Schreck, T., Preiner, R., “Semi-automated Annotation of Repetitive Ornaments on 3D Painted Pottery Surfaces”, in *Eurographics Workshop on Graphics and Cultural Heritage*, The Eurographics Association, 2020, pp. 1–4. DOI: [10.2312/gch.20201285](#) (cit. on pp. [8](#), [9](#), [12–14](#), [98](#), [100](#), [168](#), [182](#)).
- [82] Lengauer, S., Sipiran, I., Preiner, R., Schreck, T., Bustos, B., “A Benchmark Dataset for Repetitive Pattern Recognition on Textured 3D Surfaces”, *Computer Graphics Forum*, 2021. DOI: [10.1111/cgf.14352](#) (cit. on pp. [8](#), [9](#), [12–14](#), [98](#), [100](#), [181](#)).

- [83] Lengauer, S., Houska, P., Preiner, R., Trinkl, E., Karl, S., Sipiran, I., Bustos, B., Schreck, T., “Interactive annotation of geometric ornamentation on painted pottery assisted by deep learning”, *it - Information Technology*, 2022. DOI: [doi:10.1515/itit-2022-0007](https://doi.org/10.1515/itit-2022-0007) (cit. on pp. 8, 9, 12–14, 98, 100, 181).
- [84] Engel, D., Hüttenberger, L., Hamann, B., “A survey of dimension reduction methods for high-dimensional data analysis and visualization”, in *Visualization of Large and Unstructured Data Sets: Applications in Geospatial Planning, Modeling and Engineering- Proceedings of IRTG 1131 Workshop 2011*, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2012 (cit. on p. 8).
- [85] Schreck, T., Lengauer, S., Preiner, R., “Forschungsprojekt crossmodale suche und visuelle exploration von 3d kulturgutobjekten”, in *Forum Archaeologiae*, vol. 91, 2019 (cit. on pp. 9, 12).
- [86] Trinkl, E., Karl, S., Lengauer, S., Preiner, R., Schreck, T., “Cross-modal search and exploration of greek painted pottery”, in *The 3 Dimensions of Digitalised Archaeology. State-of-the-art, Data Management and Current Challenges in Archaeological 3D Documentation* (cit. on pp. 9, 12, 13).
- [87] Bustos, B., Keim, D. A., Saupe, D., Schreck, T., Vranić, D. V., “Feature-based similarity search in 3d object databases”, *ACM Computing Surveys (CSUR)*, vol. 37, no. 4, pp. 345–387, 2005 (cit. on pp. 10, 29, 43, 44, 64).
- [88] Gregor, R., Mayrbrugger, C., Mavridis, P., Bustos, B., Schreck, T., “Cross-modal content-based retrieval for digitized 2d and 3d cultural heritage artifacts”, in *Proceedings of the Eurographics Workshop on Graphics and Cultural Heritage*, Eurographics Association, 2017, pp. 119–123 (cit. on pp. 10, 46, 47, 50, 53, 57).
- [89] Tyukin, I., Sofeikov, K., Levesley, J., Gorban, A. N., Allison, P., Cooper, N. J., “Exploring automated pottery identification [arch-i-scan]”, *Internet Archaeology*, vol. 50, no. 10.11141, 2018 (cit. on pp. 10, 46, 47, 50).
- [90] Núñez Jareño, S. J., Helden, D. P., Mirkes, E. M., Tyukin, I. Y., Allison, P. M., “Learning from scarce information: Using synthetic data to classify roman fine ware pottery”, *Entropy*, vol. 23, no. 9, p. 1140, 2021 (cit. on pp. 10, 46, 47, 50).
- [91] Banterle, F., Itkin, B., Dellepiane, M., Wolf, L., Callieri, M., Dershowitz, N., Scopigno, R., “Vasesketch: Automatic 3d representation of pottery from paper catalog drawings”, in *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, IEEE, vol. 1, 2017, pp. 683–690 (cit. on pp. 10, 11, 46, 47, 51, 72).
- [92] Anichini, F., Dershowitz, N., Dubbini, N., Gattiglia, G., Itkin, B., Wolf, L., “The automatic recognition of ceramics from only one photo: The archaide app”, *Journal of Archaeological Science: Reports*, vol. 36, p. 102788, 2021 (cit. on pp. 10, 46, 47, 51).

- [93] Itkin, B., Wolf, L., Dershowitz, N., “Computational visual ceramicology: Matching image outlines to catalog sketches”, in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, 2021, pp. 14 822–14 830 (cit. on pp. [10](#), [46](#), [47](#), [51](#)).
- [94] Dalal, N., Triggs, B., “Histograms of oriented gradients for human detection”, in *CVPR (1)*, 2005, pp. 886–893 (cit. on pp. [10](#), [62](#)).
- [95] Lowe, D. G., “Distinctive image features from scale-invariant keypoints”, *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004 (cit. on pp. [10](#), [49](#), [62](#)).
- [96] Lux, M., Marques, O., “Visual information retrieval using java and lire”, *Synthesis Lectures on Information Concepts, Retrieval, and Services*, vol. 5, no. 1, pp. 1–112, 2013 (cit. on pp. [10](#), [62](#)).
- [97] Tamura, H., Mori, S., Yamawaki, T., “Textural features corresponding to visual perception”, *IEEE Transactions on Systems, man, and cybernetics*, vol. 8, no. 6, pp. 460–473, 1978 (cit. on pp. [10](#), [62](#)).
- [98] Attalla, E., Siy, P., “Robust shape similarity retrieval based on contour segmentation polygonal multiresolution and elastic matching”, *Pattern Recognition*, vol. 38, no. 12, pp. 2229–2241, 2005 (cit. on pp. [10](#), [62](#), [63](#)).
- [99] Belongie, S., Malik, J., Puzicha, J., “Shape matching and object recognition using shape contexts”, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 4, pp. 509–522, 2002 (cit. on pp. [10](#), [11](#), [87](#), [88](#), [93](#)).
- [100] He, K., Zhang, X., Ren, S., Sun, J., “Deep residual learning for image recognition”, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778 (cit. on pp. [11](#), [62](#), [77](#)).
- [101] Kirillov, A., He, K., Girshick, R., Rother, C., Dollár, P., “Panoptic segmentation”, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 9404–9413 (cit. on pp. [11](#), [31](#)).
- [102] Gonzalez, R. C., Woods, R. E., *Digital Image Processing (3rd Edition)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006, ISBN: 013168728X (cit. on pp. [11](#), [31](#), [119](#), [127](#)).
- [103] Biasotti, S., Thompson, E. M., Aono, M., Hamza, A. B., Bustos, B., Dong, S., Du, B., Fehri, A., Li, H., Limberger, F. A., Masoumi, M., Rezaei, M., Sipiran, I., Sun, L., Tatsuma, A., Forero, S. V., Wilson, R. C., Wu, Y., Zhang, J., Zhao, T., Fornasa, F., Giachetti, A., “Retrieval of surfaces with similar relief patterns: Shrec’17 track”, in *Proceedings of the Workshop on 3D Object Retrieval*, ser. 3Dor ’17, Lyon, France: Eurographics Association, 2017, pp. 95–103. DOI: [10.2312/3dor.20171058](#) (cit. on pp. [11](#), [102](#), [104](#)).

- [104] Biasotti, S., Thompson, E. M., Barthe, L., Berretti, S., Giachetti, A., Lejembre, T., Mellado, N., Moustakas, K., Manolas, I., Dimou, D., Tortorici, C., Velasco-Forero, S., Werghi, N., Polig, M., Sorrentino, G., Hermon, S., “Recognition of geometric patterns over 3d models”, in *Proceedings of the 11th Eurographics Workshop on 3D Object Retrieval*, ser. 3DOR ’18, Delft, The Netherlands: Eurographics Association, 2018, pp. 71–77. DOI: [10.2312/3dor.20181055](https://doi.org/10.2312/3dor.20181055) (cit. on pp. [11](#), [102](#), [104](#)).
- [105] Moscoso Thompson, E., Tortorici, C., Werghi, N., Berretti, S., Velasco-Forero, S., Biasotti, S., “Retrieval of Gray Patterns Depicted on 3D Models”, *Eurographics Workshop on 3D Object Retrieval*, p. 7, 2018. DOI: [10.2312/3DOR.20181054](https://doi.org/10.2312/3DOR.20181054) (cit. on pp. [11](#), [103](#), [104](#)).
- [106] Moscoso Thompson, E., Biasotti, S., Giachetti, A., Tortorici, C., Werghi, N., Obeid, A. S., Berretti, S., Nguyen-Dinh, H.-P., Le, M.-Q., Nguyen, H.-D., Tran, M.-T., Gigli, L., Velasco-Forero, S., Marcotegui, B., Sipiran, I., Bustos, B., Romanelis, I., Fotis, V., Arvanitis, G., Moustakas, K., Otu, E., Zwiggelaar, R., Hunter, D., Liu, Y., Arteaga, Y., Luxman, R., “Shrec 2020: Retrieval of digital surfaces with similar geometric reliefs”, *Computers & Graphics*, vol. 91, pp. 199–218, 2020. DOI: [10.1016/j.cag.2020.07.011](https://doi.org/10.1016/j.cag.2020.07.011) (cit. on pp. [11](#), [103](#), [104](#)).
- [107] Windhager, F., Federico, P., Schreder, G., Glinka, K., Dörk, M., Miksch, S., Mayr, E., “Visualization of cultural heritage collection data: State of the art and future challenges”, *IEEE transactions on visualization and computer graphics*, vol. 25, no. 6, pp. 2311–2330, 2018 (cit. on pp. [11](#), [20](#), [142](#), [143](#), [146](#), [147](#)).
- [108] Miksch, S., Aigner, W., “A matter of time: Applying a data–users–tasks design triangle to visual analytics of time-oriented data”, *Computers & Graphics*, vol. 38, pp. 286–290, 2014 (cit. on p. [11](#)).
- [109] Royce, W. W., “Managing the development of large software systems: Concepts and techniques”, in *Proceedings of the 9th international conference on Software Engineering*, 1987, pp. 328–338 (cit. on p. [12](#)).
- [110] (2022). Athenian agora excavations, [Online]. Available: <http://www.agathe.gr/> (cit. on p. [19](#)).
- [111] (2022). Corinth excavations - ascsa.net, [Online]. Available: <https://corinth.ascsa.net/> (cit. on p. [19](#)).
- [112] (2022). British museum, [Online]. Available: <https://www.britishmuseum.org/research.aspx> (cit. on p. [19](#)).
- [113] (2022). Collections search – museum of fine arts, boston, [Online]. Available: <https://www.mfa.org/collections/search> (cit. on p. [19](#)).
- [114] (2022). Ure museum of greek archaeology, [Online]. Available: <https://collections.reading.ac.uk/ure-museum/> (cit. on p. [19](#)).

- [115] (2022). The j. paul getty museum collection, [Online]. Available: <https://www.getty.edu/art/collection/> (cit. on p. 19).
- [116] (2022). Ashmolean, [Online]. Available: <https://collections.ashmolean.org/> (cit. on p. 19).
- [117] (2022). Digital archaeological archive of comparative slavery, [Online]. Available: <https://daacsrc.org/> (cit. on p. 19).
- [118] (2022). Europeana, [Online]. Available: <https://www.europeana.eu/> (cit. on p. 19).
- [119] (2022). 3d kulturdatenbank, landessammlung niederösterreich, [Online]. Available: <https://kulturdatenbank.at/> (cit. on pp. 19, 24).
- [120] (2022). The virtual hampson museum, [Online]. Available: <http://hampson.cast.uark.edu/> (cit. on pp. 19, 52).
- [121] (2022). Sketchfab, [Online]. Available: <https://sketchfab.com/> (cit. on p. 19).
- [122] Hafner, G., *Corpus Vasorum Antiquorum Karlsruhe, Badisches Landesmuseum 1, Deutschland 7*. München: C.H. Beck, 1951 (cit. on p. 20).
- [123] Lullies, R., *Corpus Vasorum Antiquorum München, Museum Antiker Kleinkunst 3, Deutschland 9*. München: C.H. Beck, 1952 (cit. on p. 20).
- [124] Lullies, R., *Corpus Vasorum Antiquorum Kassel, Antikenabteilung der Staatlichen Kunstsammlungen 1, Deutschland 35*. München: C.H. Beck, 1972 (cit. on p. 20).
- [125] Kauffmann-Samaras, A., *Corpus Vasorum Antiquorum Musée du Louvre 16, France 25*. Paris: E. de Boccard, 1972 (cit. on p. 20).
- [126] Hölscher, F., *Corpus Vasorum Antiquorum Würzburg, Martin von Wagner Museum 1, Deutschland 39*. München: C.H. Beck, 1975 (cit. on p. 20).
- [127] Wallenstein, K., *Corpus Vasorum Antiquorum Tübingen, Antikensammlung des Archäologischen Instituts der Universität 2, Deutschland 44*. München: C.H. Beck, 1979 (cit. on p. 20).
- [128] Kourou, N., *Corpus Vasorum Antiquorum Athens, National Museum 5, Greece 8*. Athens: Academy of Athens, 2002 (cit. on p. 20).
- [129] Moore, M., *Corpus Vasorum Antiquorum Metropolitan Museum of Art 5, U.S.A. 37*. New York: Metropolitan Museum of Art, 2004 (cit. on p. 20).
- [130] Dräger, O., *Corpus Vasorum Antiquorum Erlangen 2, Deutschland 84*. München: C.H. Beck, 2007 (cit. on p. 20).
- [131] Eschbach, N., *Corpus Vasorum Antiquorum Göttingen 3, Deutschland 83*. München: C.H. Beck, 2007 (cit. on p. 20).

- [132] Dehl-von Kaenel, C., *Corpus Vasorum Antiquorum Berlin, Antikensammlung 10, Deutschland 85*. München: C.H. Beck, 2009 (cit. on p. 20).
- [133] Pfisterer-Haas, S., *Corpus Vasorum Antiquorum Berlin, München 16, Deutschland 88*. München: C.H. Beck, 2010 (cit. on p. 20).
- [134] Catling, H., Mannack, T., *Corpus Vasorum Antiquorum Oxford, Ashmolean Museum 4, Great Britain 24*. London: Oxford University Press, 2010 (cit. on pp. 20, 27).
- [135] Coldstream, J., *Corpus Vasorum Antiquorum British Museum 11, Great Britain 25*. London: British Museum Press, 2010 (cit. on p. 20).
- [136] Schörner, H., *Corpus Vasorum Antiquorum Jena 1, Deutschland 90*. München: C.H. Beck, 2011 (cit. on p. 20).
- [137] Zimmermann-Elseify, N., *Corpus Vasorum Antiquorum Berlin, Antikensammlung 13, Deutschland 93*. München: C.H. Beck, 2013 (cit. on pp. 20, 23, 64, 89).
- [138] Hofstetter-Dolega, E., *Corpus Vasorum Antiquorum Dresden, Staatliche Kunstsammlungen, Skulpturensammlung 2, Deutschland 97*. München: C.H. Beck, 2015 (cit. on p. 20).
- [139] Piekarski, D., *Corpus Vasorum Antiquorum Bonn, Akademisches Kunstmuseum 4, Deutschland 100*. München: C.H. Beck, 2016 (cit. on p. 20).
- [140] Zimmermann-Elseify, N., *Corpus Vasorum Antiquorum Berlin, Antikensammlung 17, Deutschland 102*. München: C.H. Beck, 2018 (cit. on pp. 20, 64).
- [141] Orton, C., Hughes, M., Hughes, M., *Pottery in archaeology*. Cambridge University Press, 2013 (cit. on pp. 21, 24, 144).
- [142] Walter, C., “Towards a more ‘scientific’ archaeological tool. the accurate drawing of greek vases between the end of the nineteenth and the first half of the twentieth centuries”, *Archives, Ancestors, Practices: Archaeology in the Light of its History*. New York: Berghahn Books, pp. 179–190, 2008 (cit. on pp. 22, 30).
- [143] Westoby, M., Brasington, J., Glasser, N., Hambrey, M., Reynolds, J., “‘structure-from-motion’ photogrammetry: A low-cost, effective tool for geoscience applications”, *Geomorphology*, vol. 179, pp. 300–314, 2012. DOI: [10.1016/j.geomorph.2012.08.021](https://doi.org/10.1016/j.geomorph.2012.08.021) (cit. on pp. 24, 58).
- [144] Rieke-Zapp, D., Royo, S., “Structured light 3D scanning”, in *Digital Techniques for Documenting and Preserving Cultural Heritage*, Leeds: Arc Humanities Press, 2017, pp. 247–251 (cit. on p. 24).
- [145] Campbell, R. J., Flynn, P. J., “A survey of free-form object representation and recognition techniques”, *Computer Vision and Image Understanding*, vol. 81, no. 2, pp. 166–210, 2001 (cit. on p. 24).

- [146] Gezawa, A. S., Zhang, Y., Wang, Q., Yunqi, L., “A review on deep learning approaches for 3d data representations in retrieval and classifications”, *IEEE access*, vol. 8, pp. 57 566–57 593, 2020 (cit. on pp. [24](#), [45](#)).
- [147] Schmidt, S., “Cva today. some remarks”, in *Les cratères à volutes, Actes du Colloque de l’Académie des Inscriptions et Belles-Lettres, Paris, 8 juin 2012*, Paris: Académie des Inscriptions et Belles-Lettres, 2014, pp. 9–12 (cit. on pp. [24](#), [33](#), [58](#), [62](#)).
- [148] Unnikrishnan, R., Smith, R., “Combined script and page orientation estimation using the tesseract ocr engine”, in *Proceedings of the international workshop on multilingual OCR*, 2009, pp. 1–7 (cit. on pp. [26](#), [28](#)).
- [149] Bolelli, F., Allegretti, S., Baraldi, L., Grana, C., “Spaghetti labeling: Directed acyclic graphs for block-based connected components labeling”, *IEEE Transactions on Image Processing*, vol. 29, pp. 1999–2012, 2019 (cit. on pp. [27](#), [36](#)).
- [150] Floater, M. S., Hormann, K., “Surface parameterization: A tutorial and survey”, *Advances in multiresolution for geometric modelling*, pp. 157–186, 2005 (cit. on p. [30](#)).
- [151] Sheffer, A., Praun, E., Rose, K., “Mesh parameterization methods and their applications”, *Foundations and Trends® in Computer Graphics and Vision*, vol. 2, no. 2, pp. 105–171, 2007 (cit. on p. [30](#)).
- [152] Mara, H., Portl, J., *Acquisition and documentation of vessels using high-resolution 3D-scanners*. JSTOR, 2013 (cit. on p. [30](#)).
- [153] Rieck, B., Mara, H., Krömker, S., “Unwrapping highly-detailed 3d meshes of rotationally symmetric man-made objects”, *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 259–264, 2013. DOI: [10.11588/heidok.00015488](#) (cit. on pp. [30](#), [168](#)).
- [154] Wang, Z., Li, C., Kinoshita, T., Zhang, Z., Konno, K., “Unfoldment of surface pattern with highlighting decoration via rotationally symmetric jomon earthenware”, in *2019 Nicograph International (NicoInt)*, 2019, pp. 33–36. DOI: [10.1109/NICOInt.2019.00014](#) (cit. on p. [30](#)).
- [155] Girshick, R., Donahue, J., Darrell, T., Malik, J., “Rich feature hierarchies for accurate object detection and semantic segmentation”, in *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, ser. CVPR ’14, Washington, DC, USA: IEEE Computer Society, 2014, pp. 580–587. DOI: [10.1109/CVPR.2014.81](#) (cit. on p. [31](#)).
- [156] Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., LeCun, Y., “Overfeat: Integrated recognition, localization and detection using convolutional networks”, in *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*, 2014 (cit. on p. [31](#)).

- [157] Long, J., Shelhamer, E., Darrell, T., “Fully convolutional networks for semantic segmentation”, in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2015 (cit. on p. 31).
- [158] Ronneberger, O., Fischer, P., Brox, T., “U-net: Convolutional networks for biomedical image segmentation”, in *International Conference on Medical image computing and computer-assisted intervention*, Springer, 2015, pp. 234–241 (cit. on p. 31).
- [159] He, K., Gkioxari, G., Dollár, P., Girshick, R., “Mask r-cnn”, in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969 (cit. on p. 31).
- [160] Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S., “Feature pyramid networks for object detection”, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2117–2125 (cit. on p. 31).
- [161] Chen, X., Golovinskiy, A., Funkhouser, T., “A benchmark for 3D mesh segmentation”, *ACM Transactions on Graphics (Proc. SIGGRAPH)*, vol. 28, no. 3, Aug. 2009 (cit. on p. 31).
- [162] Sipiran, I., Bustos, B., Schreck, T., “Data-aware 3D partitioning for generic shape retrieval”, *Computers & Graphics Special Issue on 3D Object Retrieval*, vol. 37, no. 5, pp. 460–472, Aug. 2013. DOI: [10.1016/j.cag.2013.04.002](https://doi.org/10.1016/j.cag.2013.04.002) (cit. on p. 31).
- [163] Tagliasacchi, A., Delame, T., Spagnuolo, M., Amenta, N., Telea, A., “3D Skeletons: A State-of-the-Art Report”, *Computer Graphics Forum*, 2016. DOI: [10.1111/cgf.12865](https://doi.org/10.1111/cgf.12865) (cit. on p. 31).
- [164] Rother, C., Kolmogorov, V., Blake, A., “Grabcut: Interactive foreground extraction using iterated graph cuts”, *ACM Trans. Graph.*, vol. 23, no. 3, pp. 309–314, Aug. 2004. DOI: [10.1145/1015706.1015720](https://doi.org/10.1145/1015706.1015720) (cit. on pp. 31, 32, 35, 89).
- [165] Felzenszwalb, P. F., Huttenlocher, D. P., “Efficient graph-based image segmentation”, *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167–181, Sep. 2004. DOI: [10.1023/B:VISI.0000022288.19776.77](https://doi.org/10.1023/B:VISI.0000022288.19776.77) (cit. on pp. 31, 35).
- [166] Chambolle, A., Pock, T., “A first-order primal-dual algorithm for convex problems with applications to imaging”, *Journal of mathematical imaging and vision*, vol. 40, no. 1, pp. 120–145, 2011 (cit. on p. 34).
- [167] Otsu, N., “A threshold selection method from gray-level histograms”, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, Jan. 1979. DOI: [10.1109/TSMC.1979.4310076](https://doi.org/10.1109/TSMC.1979.4310076) (cit. on pp. 34, 120).
- [168] Karasik, A., Smilansky, U., “3d scanning technology as a standard archaeological tool for pottery analysis: Practice and theory”, *Journal of Archaeological Science*, vol. 35, no. 5, pp. 1148–1168, 2008. DOI: [10.1016/j.jas.2007.08.008](https://doi.org/10.1016/j.jas.2007.08.008) (cit. on pp. 42, 43).
- [169] Kunze, E., “Disiecta membra attischer grabkratere”, *AEphem*, pp. 162–171, 1953 (cit. on p. 42).

- [170] Baeza-Yates, R., Ribeiro-Neto, B., *Modern information retrieval*. ACM press New York, 1999, vol. 463 (cit. on pp. 43, 64).
- [171] Gudivada, V. N., Raghavan, V. V., “Content based image retrieval systems”, *Computer*, vol. 28, no. 9, pp. 18–22, 1995 (cit. on p. 43).
- [172] Latif, A., Rasheed, A., Sajid, U., Ahmed, J., Ali, N., Ratyal, N. I., Zafar, B., Dar, S. H., Sajid, M., Khalil, T., “Content-based image retrieval and feature extraction: A comprehensive review”, *Mathematical Problems in Engineering*, vol. 2019, 2019 (cit. on p. 44).
- [173] Tangelder, J., Veltkamp, R., “A survey of content based 3D shape retrieval methods”, *Multimedia Tools and Applications*, vol. 39, no. 3, pp. 441–471, 2008. DOI: [10.1007/s11042-007-0181-0](https://doi.org/10.1007/s11042-007-0181-0) (cit. on p. 44).
- [174] Schreck, T., “What features can tell us about shape”, *IEEE Computer Graphics and Applications*, vol. 37, no. 3, pp. 82–87, 2017. DOI: [10.1109/MCG.2017.41](https://doi.org/10.1109/MCG.2017.41) (cit. on p. 44).
- [175] Bai, S., Bai, X., Zhou, Z., Zhang, Z., Tian, Q., Latecki, L. J., “GIFT: Towards scalable 3D shape retrieval”, *IEEE Trans. Multimedia*, vol. 19, no. 6, pp. 1257–1271, 2017. DOI: [10.1109/TMM.2017.2652071](https://doi.org/10.1109/TMM.2017.2652071) (cit. on p. 45).
- [176] Ioannidou, A., Chatzilari, E., Nikolopoulos, S., Kompatsiaris, I., “Deep learning advances in computer vision with 3d data: A survey”, *ACM Comput. Surv.*, vol. 50, no. 2, 20:1–20:38, 2017. DOI: [10.1145/3042064](https://doi.org/10.1145/3042064) (cit. on p. 45).
- [177] Savelonas, M., Pratikakis, I., Sfikas, K., “An overview of partial 3D object retrieval methodologies”, *Multimedia Tools and Applications*, pp. 1–26, 2014 (cit. on p. 45).
- [178] Lavoué, G., “Combination of bag-of-words descriptors for robust partial shape retrieval”, *The Visual Computer*, vol. 28, no. 9, pp. 931–942, 2012 (cit. on p. 45).
- [179] Bronstein, A. M., Bronstein, M. M., Guibas, L. J., Ovsjanikov, M., “Shape google: Geometric words and expressions for invariant shape retrieval”, *ACM Transactions on Graphics (TOG)*, vol. 30, no. 1, pp. 1–20, 2011 (cit. on p. 45).
- [180] Sipiran, I., Meruane, R., Bustos, B., Schreck, T., Li, B., Lu, Y., Johan, H., “A benchmark of simulated range images for partial shape retrieval”, *The Visual Computer*, vol. 30, no. 11, pp. 1293–1308, 2014 (cit. on p. 45).
- [181] Grünewald, M., “Ausgrabungen im legionslager von carnuntum (grabungen 1969-1977). keramik und kleinfunde 1976-1977”, *Der römische Limes in Österreich*, vol. 34, pp. 10–11, 1986 (cit. on pp. 46–48).
- [182] Rotroff, S. I., “Hellenistic pottery: Athenian and imported wheelmade table ware and related material (the athenian agora 29)”, *Princeton (NJ): American School of Classical Studies at Athens*, 1997 (cit. on p. 46).
- [183] Rotroff, S. I., *Hellenistic pottery: the plain wares*. ASCSA, 2006, vol. 33 (cit. on p. 46).

- [184] Hayes, J., “Late roman pottery. london”, *British School at Rome*, 1972 (cit. on p. 46).
- [185] Bishop, G., Cha, S.-H., Tappert, C., “A greek pottery shape and school identification and classification system using image retrieval techniques”, *Proceedings of Student / Faculty Research Day CSIS. Pace University*, 2005 (cit. on pp. 46, 47, 50).
- [186] Biasotti, S., Cerri, A., Falcidieno, B., Spagnuolo, M., “3d artifacts similarity based on the concurrent evaluation of heterogeneous properties”, *Journal on Computing and Cultural Heritage (JOCCH)*, vol. 8, no. 4, pp. 1–19, 2015 (cit. on pp. 47, 52).
- [187] Stergioulas, A., Ioannakis, G., Koutsoudis, A., Chamzas, C., “The orion pottery repository – a publicly available 3d objects’ benchmark database with texture information”, in *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection*, Cham: Springer International Publishing, 2018, pp. 174–187, ISBN: 978-3-030-01762-0 (cit. on pp. 47, 52).
- [188] Kampel, M., Sablatnig, R., “Automated segmentation of archaeological profiles for classification”, in *2002 International Conference on Pattern Recognition*, IEEE, vol. 1, 2002, pp. 57–60 (cit. on p. 48).
- [189] Birkhoff, G. D., “Aesthetic measure”, in *Aesthetic Measure*, Harvard University Press, 2013 (cit. on p. 49).
- [190] Karasik, A., Smilansky, U., “Computerized morphological classification of ceramics”, *Journal of Archaeological Science*, vol. 38, no. 10, pp. 2644–2657, 2011. DOI: [10.1016/j.jas.2011.05.023](https://doi.org/10.1016/j.jas.2011.05.023) (cit. on p. 50).
- [191] Navarro, P., Cintas, C., Lucena, M., Fuertes, J. M., Delrieux, C., Molinos, M., “Learning feature representation of iberian ceramics with automatic classification models”, *Journal of Cultural Heritage*, vol. 48, pp. 65–73, 2021. DOI: [10.1016/j.culher.2021.01.003](https://doi.org/10.1016/j.culher.2021.01.003) (cit. on p. 50).
- [192] Parisotto, S., Leone, N., Schönlieb, C.-B., Launaro, A., “Unsupervised clustering of roman potsherds via variational autoencoders”, *Journal of Archaeological Science*, vol. 142, p. 105 598, 2022. DOI: [10.1016/j.jas.2022.105598](https://doi.org/10.1016/j.jas.2022.105598) (cit. on p. 50).
- [193] Darom, T., Keller, Y., “Scale-invariant features for 3-d mesh models”, *IEEE Transactions on Image Processing*, vol. 21, no. 5, pp. 2758–2769, 2012 (cit. on p. 52).
- [194] Perronnin, F., Dance, C., “Fisher kernels on visual vocabularies for image categorization”, in *2007 IEEE conference on computer vision and pattern recognition*, IEEE, 2007, pp. 1–8 (cit. on p. 52).
- [195] Willis, A., Orriols, X., Cooper, D. B., “Accurately estimating sherd 3d surface geometry with application to pot reconstruction”, in *2003 Conference on Computer Vision and Pattern Recognition Workshop*, IEEE, vol. 1, 2003, pp. 5–5 (cit. on p. 59).

- [196] Cao, Y., Mumford, D., “Geometric structure estimation of axially symmetric pots from small fragments”, in *Signal Processing, Pattern Recognition, and Applications, IASTED International Conference*, Citeseer, vol. 2, 2002 (cit. on p. 59).
- [197] Han, D., Hahn, H.-S., “Axis estimation and grouping of rotationally symmetric object segments”, *Pattern recognition*, vol. 47, no. 1, pp. 296–312, 2014 (cit. on p. 59).
- [198] Di Angelo, L., Di Stefano, P., “Axis estimation of thin-walled axially symmetric solids”, *Pattern Recognition Letters*, vol. 106, pp. 47–52, 2018 (cit. on p. 59).
- [199] Hong, J. H., Kim, Y. M., Wi, K.-C., Kim, J., “Potsac: A robust axis estimator for axially symmetric pot fragments.”, in *ICCV Workshops*, 2019, pp. 1421–1428 (cit. on p. 59).
- [200] Hays, J., Efros, A. A., “Scene completion using millions of photographs”, *ACM Transactions on Graphics (ToG)*, vol. 26, no. 3, 4-es, 2007 (cit. on p. 60).
- [201] Whyte, O., Sivic, J., Zisserman, A., “Get out of my picture! internet-based inpainting.”, in *BMVC*, vol. 2, 2009, p. 5 (cit. on p. 60).
- [202] Jam, J., Kendrick, C., Walker, K., Drouard, V., Hsu, J. G.-S., Yap, M. H., “A comprehensive review of past and present image inpainting methods”, *Computer Vision and Image Understanding*, vol. 203, p. 103 147, 2021. DOI: [10.1016/j.cviu.2020.103147](https://doi.org/10.1016/j.cviu.2020.103147) (cit. on pp. 60, 61).
- [203] Wei, L.-Y., Levoy, M., “Fast texture synthesis using tree-structured vector quantization”, in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, 2000, pp. 479–488 (cit. on pp. 60, 61).
- [204] Lazebnik, S., Schmid, C., Ponce, J., “A sparse texture representation using local affine regions”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1265–1278, 2005. DOI: [10.1109/TPAMI.2005.151](https://doi.org/10.1109/TPAMI.2005.151) (cit. on p. 60).
- [205] Criminisi, A., Pérez, P., Toyama, K., “Region filling and object removal by exemplar-based image inpainting”, *IEEE Transactions on image processing*, vol. 13, no. 9, pp. 1200–1212, 2004 (cit. on pp. 61, 162, 163).
- [206] Seddati, O., Dupont, S., Mahmoudi, S., Parian, M., “Towards good practices for image retrieval based on cnn features”, in *Proceedings of the IEEE international conference on computer vision workshops*, 2017, pp. 1246–1255 (cit. on p. 62).
- [207] Suzuki, S., “Topological structural analysis of digitized binary images by border following”, *Computer vision, graphics, and image processing*, vol. 30, no. 1, pp. 32–46, 1985 (cit. on pp. 63, 148).
- [208] Henderson, C., Izquierdo, E., “Symmetric stability of low level feature detectors”, *Pattern Recognition Letters*, vol. 78, pp. 36–40, 2016. DOI: [10.1016/j.patrec.2016.03.027](https://doi.org/10.1016/j.patrec.2016.03.027) (cit. on p. 64).

- [209] Kanezaki, A., Mukuta, Y., Harada, T., “Mirror reflection invariant hog descriptors for object detection”, in *2014 IEEE International Conference on Image Processing (ICIP)*, IEEE, 2014, pp. 1594–1598 (cit. on p. 64).
- [210] Zhou, X. S., Huang, T. S., “Relevance feedback in image retrieval: A comprehensive review”, *Multimedia systems*, vol. 8, no. 6, pp. 536–544, 2003 (cit. on p. 64).
- [211] Elad, M., Tal, A., Ar, S., “Content based retrieval of vrml objects—an iterative and interactive approach”, in *Multimedia 2001*, Springer, 2002, pp. 107–118 (cit. on p. 64).
- [212] Rocchio, J., “Relevance feedback in information retrieval”, *The Smart retrieval system-experiments in automatic document processing*, pp. 313–323, 1971 (cit. on p. 64).
- [213] Haspels, C., *Attic Black-figured Lekythoi*, ser. Travaux et mémoires / École française d’Athènes 4. “de” Boccard, 1936 (cit. on p. 70).
- [214] Lang-Auinger, C., Karl, S., Kratzmüller, B., *Corpus Vasorum Antiquorum Wien, Kunsthistorisches Museum 6, Österreich 7*. Vienna: Verlag der Österreichischen Akademie der Wissenschaften, 2019 (cit. on p. 71).
- [215] Gregor, R., Bauer, D., Sipiran, I., Perakis, P., Schreck, T., “Automatic 3d object fracturing for evaluation of partial retrieval and object restoration tasks : Benchmark and application to 3d cultural heritage data”, in *Eurographics Workshop on 3D Object Retrieval*, Goslar: The Eurographics Association, 2015, pp. 7–14. DOI: [10.2312/3dor.20151049](https://doi.org/10.2312/3dor.20151049) (cit. on p. 71).
- [216] Mokhtarian, F., Abbasi, S., “Shape similarity retrieval under affine transforms”, *Pattern Recognition*, vol. 35, no. 1, pp. 31–41, 2002 (cit. on p. 73).
- [217] Rosenfeld, A., Johnston, E., “Angle detection on digital curves”, *IEEE Transactions on Computers*, vol. 100, no. 9, pp. 875–878, 1973 (cit. on p. 73).
- [218] Zeng, L., Liu, Y.-J., Wang, J., Zhang, D.-L., Yuen, M. M.-F., “Sketch2jewelry: Semantic feature modeling for sketch-based jewelry design”, *Computers & Graphics*, vol. 38, pp. 69–77, Feb. 2014. DOI: [10.1016/j.cag.2013.10.017](https://doi.org/10.1016/j.cag.2013.10.017) (cit. on p. 78).
- [219] Itskovich, A., Tal, A., “Surface partial matching and application to archaeology”, *Computers & Graphics*, vol. 35, no. 2, pp. 334–341, 2011, Virtual Reality in Brazil Visual Computing in Biology and Medicine Semantic 3D media and content Cultural Heritage. DOI: [10.1016/j.cag.2010.11.010](https://doi.org/10.1016/j.cag.2010.11.010) (cit. on pp. 85, 86, 103).
- [220] Leifman, G., Tal, A., “Pattern-Driven Colorization of 3D Surfaces”, in *2013 IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 241–248 (cit. on p. 85).
- [221] Crowley, E. J., Zisserman, A., “Of gods and goats: Weakly supervised learning of figurative art”, *learning*, vol. 8, p. 14, 2013 (cit. on pp. 85, 183).

- [222] Kipke, M., Brinkmeyer, L., Bagayoko, S., Schmidt-Thieme, L., Langner, M., “Deep level annotation for painter attribution on greek vases utilizing object detection”, in *Proceedings of the 4th ACM International Workshop on Structuring and Understanding of Multimedia HeritAge Contents*, ser. SUMAC '22, Lisboa, Portugal: Association for Computing Machinery, 2022, pp. 23–31. DOI: [10.1145/3552464.3555684](https://doi.org/10.1145/3552464.3555684) (cit. on pp. [85](#), [86](#), [137](#)).
- [223] Bendschus, T., Madhu, P., “Deep learning based attribute representation in ancient vase paintings”, in *DH2020* (cit. on p. [86](#)).
- [224] Kuhn, H. W., “The hungarian method for the assignment problem”, *Naval Research Logistics Quarterly*, vol. 2, no. 1-2, pp. 83–97, 1955. DOI: [10.1002/nav.3800020109](https://doi.org/10.1002/nav.3800020109) (cit. on p. [88](#)).
- [225] Jonker, R., Volgenant, A., “A shortest augmenting path algorithm for dense and sparse linear assignment problems”, *Computing*, vol. 38, no. 4, pp. 325–340, 1987 (cit. on p. [88](#)).
- [226] Bhattacharyya, A., “On a measure of divergence between two statistical populations defined by their probability distributions”, *Bull. Calcutta Math. Soc.*, vol. 35, pp. 99–109, 1943 (cit. on pp. [89](#), [124](#)).
- [227] Eschbach, N., *Corpus Vasorum Antiquorum Dresden, Staatliche Kunstsammlungen, Skulpturensammlung 3, Deutschland 104*. München: C.H. Beck, 2018 (cit. on p. [89](#)).
- [228] Dulecha, T., Giachetti, A., Pintus, R., Ciortan, I.-M., Jaspe Villanueva, A., Gobbetti, E., “Crack detection in single-and multi-light images of painted surfaces using convolutional neural networks”, in *Eurographics Workshop on Graphics and Cultural Heritage*, The Eurographics Association, 2019, pp. 43–50 (cit. on pp. [93](#), [176](#)).
- [229] Guerrero, P., Auzinger, T., Wimmer, M., Jeschke, S., “Partial shape matching using transformation parameter similarity”, *Computer Graphics Forum*, vol. 34, no. 1, pp. 239–252, 2015. DOI: [10.1111/cgf.12509](https://doi.org/10.1111/cgf.12509) (cit. on p. [93](#)).
- [230] Mayerhofer, M., “Visual exploration of repetitive patterns in cultural heritage objects”, *Master’s thesis, Graz University of Technology, Institute of Computer Graphics and Knowledge Visualization*, 2022 (cit. on pp. [99](#), [116](#), [138](#), [183](#)).
- [231] Zhou, J., Smith, K., Wilsbacher, G., Sagona, P., Reddy, D., Torkian, B., “Building science gateways for humanities”, in *Practice and Experience in Advanced Research Computing*, 2020, pp. 327–332 (cit. on p. [101](#)).
- [232] Moscoso Thompson, E., Biasotti, S., “Description and retrieval of geometric patterns on surface meshes using an edge-based LBP approach”, *Pattern Recognition*, vol. 82, pp. 1–15, 2018. DOI: [10.1016/j.patcog.2018.04.028](https://doi.org/10.1016/j.patcog.2018.04.028) (cit. on p. [101](#)).

- [233] Thompson, E. M., Biasotti, S., Digne, J., Chaine, R., “mpLBP: An Extension of the Local Binary Pattern to Surfaces based on an Efficient Coding of the Point Neighbours”, *Eurographics Workshop on 3D Object Retrieval*, 8 pages, 2019, Artwork Size: 8 pages isbn: 9783038680772 Publisher: The Eurographics Association Version Number: 009-016. DOI: [10.2312/3DOR.20191056](https://doi.org/10.2312/3DOR.20191056) (cit. on p. 101).
- [234] Bogacz, B., Mara, H., “Automatable annotations—image processing and machine learning for script in 3d and 2d with gigamesh”, *Kodikologie und Paläographie im Digitalen Zeitalter*, pp. 137–149, 2017 (cit. on p. 101).
- [235] Rodriguez-Echavarria, K., Song, R., “Analyzing the decorative style of 3d heritage collections based on shape saliency”, *JOCCH*, vol. 9, no. 4, 20:1–20:17, 2016. DOI: [10.1145/2943778](https://doi.org/10.1145/2943778) (cit. on pp. 101, 102).
- [236] Gal, R., Cohen-Or, D., “Salient Geometric Features for Partial Shape Matching and Similarity”, *ACM Trans. Graph.*, vol. 25, no. 1, pp. 130–150, Jan. 2006. DOI: [10.1145/1122501.1122507](https://doi.org/10.1145/1122501.1122507) (cit. on pp. 101, 102).
- [237] Mitra, N. J., Bronstein, A., Bronstein, M., “Intrinsic Regularity Detection in 3D Geometry”, in *ECCV 2010*, Berlin, Heidelberg, 2010, pp. 398–410, ISBN: 978-3-642-15558-1 (cit. on pp. 101, 102).
- [238] Moscoso Thompson, E., Biasotti, S., “Retrieving color patterns on surface meshes using edgeLBP descriptors”, *Computers & Graphics*, vol. 79, pp. 46–57, 2019. DOI: [10.1016/j.cag.2018.12.009](https://doi.org/10.1016/j.cag.2018.12.009) (cit. on p. 101).
- [239] Pawlowicz, L. M., Downum, C. E., “Applications of deep learning to decorated ceramic typology and classification: A case study using tusayan white ware from northeast arizona”, *Journal of Archaeological Science*, vol. 130, p. 105 375, 2021. DOI: [10.1016/j.jas.2021.105375](https://doi.org/10.1016/j.jas.2021.105375) (cit. on p. 101).
- [240] Úbeda, I., Saavedra, J. M., Nicolas, S., Petitjean, C., Heutte, L., “Improving pattern spotting in historical documents using feature pyramid networks”, *Pattern Recognition Letters*, vol. 131, pp. 398–404, 2020. DOI: [10.1016/j.patrec.2020.02.002](https://doi.org/10.1016/j.patrec.2020.02.002) (cit. on pp. 101, 102).
- [241] Werghi, N., Berretti, S., Del Bimbo, A., “The mesh-lbp: A framework for extracting local binary patterns from discrete manifolds”, *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. 220–235, 2014 (cit. on p. 101).
- [242] Gal, R., Cohen-Or, D., “Salient geometric features for partial shape matching and similarity”, *ACM Trans. Graph.*, vol. 25, no. 1, pp. 130–150, Jan. 2006. DOI: [10.1145/1122501.1122507](https://doi.org/10.1145/1122501.1122507) (cit. on p. 103).

- [243] Mitra, N. J., Bronstein, A., Bronstein, M., “Intrinsic regularity detection in 3d geometry”, in *Computer Vision – ECCV 2010*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 398–410, ISBN: 978-3-642-15558-1 (cit. on p. 103).
- [244] Berner, A., Wand, M., Mitra, N. J., Mewes, D., Seidel, H.-P., “Shape analysis with subspace symmetries”, *Computer Graphics Forum*, vol. 30, no. 2, pp. 277–286, 2011. DOI: [10.1111/j.1467-8659.2011.01859.x](#) (cit. on p. 103).
- [245] Kalojanov, J., Bokeloh, M., Wand, M., Guibas, L., Seidel, H.-P., Slusallek, P., “Microtiles: Extracting building blocks from correspondences”, *Computer Graphics Forum*, vol. 31, no. 5, pp. 1597–1606, 2012. DOI: [10.1111/j.1467-8659.2012.03165.x](#) (cit. on p. 103).
- [246] Huang, Q., Guibas, L. J., Mitra, N. J., “Near-regular structure discovery using linear programming”, *ACM Trans. Graph.*, vol. 33, no. 3, Jun. 2014. DOI: [10.1145/2535596](#) (cit. on p. 103).
- [247] Beauxis-Aussalet, E., Behrisch, M., Borgo, R., Chau, D. H., Collins, C., Ebert, D., El-Assady, M., Endert, A., Keim, D. A., Kohlhammer, J., “The role of interactive visualization in fostering trust in ai”, *IEEE Computer Graphics and Applications*, vol. 41, no. 6, pp. 7–12, 2021 (cit. on p. 103).
- [248] Endert, A., Ribarsky, W., Turkay, C., Wong, B. W., Nabney, I., Blanco, I. D., Rossi, F., “The state of the art in integrating machine learning into visual analytics”, in *Computer Graphics Forum*, Wiley Online Library, vol. 36, 2017, pp. 458–486 (cit. on p. 103).
- [249] Chegini, M., Bernard, J., Cui, J., Chegini, F., Sourin, A., Andrews, K., Schreck, T., “Interactive visual labelling versus active learning: An experimental comparison”, *Frontiers of Information Technology & Electronic Engineering*, vol. 21, no. 4, pp. 524–535, 2020 (cit. on pp. 103, 104).
- [250] Bernard, J., Zeppelzauer, M., Sedlmair, M., Aigner, W., “A unified process for visual-interactive labeling”, in *Proceedings of the EuroVis Workshop on Visual Analytics*, ser. EuroVA ’17, Barcelona, Spain: Eurographics Association, 2017, pp. 73–77. DOI: [10.2312/eurova.20171123](#) (cit. on pp. 103, 104).
- [251] Loop, C., “Smooth subdivision surfaces based on triangles”, *Master’s thesis, University of Utah, Department of Mathematics*, 1987 (cit. on p. 108).
- [252] Everingham, M., Eslami, S. M. A., Van Gool, L., Williams, C. K. I., Winn, J., Zisserman, A., “The pascal visual object classes challenge: A retrospective”, *International Journal of Computer Vision*, vol. 111, no. 1, pp. 98–136, Jan. 2015 (cit. on p. 110).
- [253] Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., Alsaadi, F. E., “A survey of deep neural network architectures and their applications”, *Neurocomputing*, vol. 234, pp. 11–26, 2017 (cit. on p. 118).

- [254] Steckner, C., in. Freiburg, 1988, ch. Begriffliche und empirische Objektordnung In: C. Steckner, Archäologie und Neue Technologien, Schriften des Deutschen Archäologen-Verbandes 10, Deutscher Archäologen-Verband, pp. 102–139 (cit. on p. 118).
- [255] Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., Bouchachia, A., “A survey on concept drift adaptation”, *ACM computing surveys (CSUR)*, vol. 46, no. 4, pp. 1–37, 2014 (cit. on p. 122).
- [256] Cook, J. M., “Kerameikos. ergebnisse der ausgrabungen. vol. v, 1. die nekropole des 10. bis 8. jahrhunderts. text and plates. by k. kübler. pp. viii 310, with 4 plans. 167 plates. berlin: De gruyter, 1954. dm. 178.”, *The Journal of Hellenic Studies*, vol. 76, pp. 124–125, 1956. DOI: [10.2307/629585](https://doi.org/10.2307/629585) (cit. on pp. 125, 126).
- [257] Bobadilla, J., Ortega, F., Hernando, A., Bernal, J., “A collaborative filtering approach to mitigate the new user cold start problem”, *Knowledge-based systems*, vol. 26, pp. 225–238, 2012 (cit. on p. 127).
- [258] Minaee, S., Boykov, Y. Y., Porikli, F., Plaza, A. J., Kehtarnavaz, N., Terzopoulos, D., “Image segmentation using deep learning: A survey”, *IEEE transactions on pattern analysis and machine intelligence*, 2021 (cit. on p. 127).
- [259] Canny, J., “A computational approach to edge detection”, *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 679–698, 1986 (cit. on p. 131).
- [260] Singh, T. R., Roy, S., Singh, O. I., Sinam, T., Singh, K. M., *A new local adaptive thresholding technique in binarization*, 2012. arXiv: [1201.5227 \[cs.CV\]](https://arxiv.org/abs/1201.5227) (cit. on p. 133).
- [261] Blackledge, J. M., *Digital image processing: mathematical and computational methods*. Elsevier, 2005 (cit. on p. 133).
- [262] Moretti, F., *Graphs, maps, trees: abstract models for a literary history*. Verso, 2005 (cit. on p. 143).
- [263] Stefan, J., Franzini, G., Cheema, M. F., Gerik, S., “On close and distant reading in digital humanities: A survey and future challenges”, in *Eurographics Conference on Visualization (EuroVis)(2015)*, R. Borgo, F. Ganovelli, I. Viola, 2015, N–A (cit. on p. 143).
- [264] Brivio, P., Tarini, M., Cignoni, P., “Browsing large image datasets through voronoi diagrams”, *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 6, pp. 1261–1270, 2010. DOI: [10.1109/TVCG.2010.136](https://doi.org/10.1109/TVCG.2010.136) (cit. on p. 143).
- [265] Glinka, K., Pietsch, C., Dörk, M., “Past visions and reconciling views: Visualizing time, texture and themes in cultural collections”, *Digit. Humanit. Q.*, vol. 11, 2017 (cit. on p. 143).
- [266] Hadlak, S., Schumann, H., Schulz, H.-J., “A Survey of Multi-faceted Graph Visualization”, in *Eurographics Conference on Visualization (EuroVis) - STARs*, The Eurographics Association, 2015. DOI: [10.2312/eurovisstar.20151109](https://doi.org/10.2312/eurovisstar.20151109) (cit. on p. 143).

- [267] Landesberger, T., Kuijper, A., Schreck, T., Kohlhammer, J., Wijk, J. J., Fekete, J., Fellner, D. W., “Visual analysis of large graphs: State-of-the-art and future research challenges”, *Comput. Graph. Forum*, vol. 30, no. 6, pp. 1719–1749, 2011. DOI: [10.1111/j.1467-8659.2011.01898.x](#) (cit. on p. [143](#)).
- [268] Tarawaneh, R. M., Keller, P., Ebert, A., “A General Introduction To Graph Visualization Techniques”, in *Visualization of Large and Unstructured Data Sets: Applications in Geospatial Planning, Modeling and Engineering - Proceedings of IRTG 1131 Workshop 2011*, ser. OpenAccess Series in Informatics (OASIs), vol. 27, Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2012, pp. 151–164. DOI: [10.4230/OASIs.VLUDS.2011.151](#) (cit. on p. [143](#)).
- [269] Roberts, J., Al-manee, H., Butcher, P., Lew, R., Rees, G., Sharma, N., Frankenberg-Garcia, A., “Multiple views: Different meanings and collocated words”, in *Computer Graphics Forum*, Wiley Online Library, vol. 38, 2019, pp. 79–93 (cit. on pp. [143](#), [145](#)).
- [270] Bavoil, L., Callahan, S. P., Scheidegger, C. E., Vo, H. T., Crossno, P., Silva, C. T., Freire, J., “Vistrails: Enabling interactive multiple-view visualizations”, in *16th IEEE Visualization Conference, VIS 2005, Minneapolis, MN, USA, October 23-28, 2005*, 2005, pp. 135–142 (cit. on p. [144](#)).
- [271] Weaver, C., “Building highly-coordinated visualizations in improvise”, in *10th IEEE Symposium on Information Visualization (InfoVis 2004), 10-12 October 2004, Austin, TX, USA*, 2004, pp. 159–166 (cit. on p. [144](#)).
- [272] Steed, C. A., Goodall, J. R., Chae, J., Trofimov, A., “Crossvis: A visual analytics system for exploring heterogeneous multivariate data with applications to materials and climate sciences”, *Graphics and Visual Computing*, p. 200 013, 2020 (cit. on p. [144](#)).
- [273] Fruchterman, T. M., Reingold, E. M., “Graph drawing by force-directed placement”, *Software: Practice and experience*, vol. 21, no. 11, pp. 1129–1164, 1991 (cit. on p. [146](#)).
- [274] Windhager, F., Mayr, E., Schreder, G., Smuc, M., Federico, P., Miksch, S., “Reframing cultural heritage collections in a visualization framework of space-time cubes.”, in *HistoInformatics@ DH*, 2016, pp. 20–24 (cit. on pp. [146](#), [148](#)).
- [275] Song, H., Fu, Y., Saket, B., Stasko, J., “Understanding the effects of visualizing missing values on visual data exploration”, in *2021 IEEE Visualization Conference (VIS)*, IEEE, 2021, pp. 161–165 (cit. on p. [147](#)).
- [276] Bernhard, K., Vygen, J., “Combinatorial optimization: Theory and algorithms”, *Springer, Third Edition, 2005.*, 2008 (cit. on p. [148](#)).
- [277] Satuluri, V., Parthasarathy, S., Ruan, Y., “Local graph sparsification for scalable clustering”, in *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, 2011, pp. 721–732 (cit. on pp. [149](#), [150](#), [182](#), [188](#), [189](#)).

-
- [278] Herman, I., Melançon, G., Marshall, M. S., “Graph visualization and navigation in information visualization: A survey”, *IEEE Transactions on visualization and computer graphics*, vol. 6, no. 1, pp. 24–43, 2000 (cit. on p. 151).
 - [279] Saunders, D., “The distribution of the berlin painter’s vases”, in *The Berlin Painter and His World: Athenian Vase-Painting in the Early Fifth Century B.C.* Princeton University Art Museum, 2017, pp. 107–131 (cit. on p. 153).
 - [280] Schmidt, S., Steinhart, M., “Sammeln und erforschen: Griechische vasen in neuzeitlichen sammlungen”, in *Beihefte zum Corpus Vasorum Antiquorum Deutschland 6*, Verlag CH Beck, Munich, 2014 (cit. on p. 154).
 - [281] Steinberger, M., Waldner, M., Streit, M., Lex, A., Schmalstieg, D., “Context-preserving visual links”, *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2249–2258, 2011 (cit. on p. 156).
 - [282] Papaioannou, G., Schreck, T., Andreadis, A., Mavridis, P., Gregor, R., Sipiran, I., Vardis, K., “From reassembly to object completion: A complete systems pipeline”, *Journal on Computing and Cultural Heritage (JOCCH)*, vol. 10, no. 2, pp. 1–22, 2017 (cit. on pp. 162, 165).
 - [283] Sipiran, I., Gregor, R., Schreck, T., “Approximate symmetry detection in partial 3d meshes”, in *Computer Graphics Forum*, Wiley Online Library, vol. 33, 2014, pp. 131–140 (cit. on p. 162).
 - [284] Gregor, R., Sipiran, I., Papaioannou, G., Schreck, T., Andreadis, A., Mavridis, P., “Towards automated 3d reconstruction of defective cultural heritage objects.”, in *GCH*, 2014, pp. 135–144 (cit. on p. 162).
 - [285] Hermoza, R., Sipiran, I., “3d reconstruction of incomplete archaeological objects using a generative adversarial network”, in *Proceedings of Computer Graphics International 2018*, ser. CGI 2018, Bintan, Island, Indonesia: Association for Computing Machinery, 2018, pp. 5–11. DOI: [10.1145/3208159.3208173](https://doi.org/10.1145/3208159.3208173) (cit. on p. 162).
 - [286] Lamb, N., Banerjee, S., Banerjee, N. K., “Automated reconstruction of smoothly joining 3d printed restorations to fix broken objects”, in *Proceedings of the ACM Symposium on Computational Fabrication*, 2019, pp. 1–12 (cit. on pp. 162, 164).
 - [287] Jboor, N. H., Belhi, A., Al-Ali, A. K., Bouras, A., Jaoua, A., “Towards an inpainting framework for visual cultural heritage”, in *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)*, IEEE, 2019, pp. 602–607 (cit. on p. 162).
 - [288] Wang, N., Wang, W., Hu, W., Fenster, A., Li, S., “Thanka mural inpainting based on multi-scale adaptive partial convolution and stroke-like mask”, *IEEE Transactions on Image Processing*, vol. 30, pp. 3720–3733, 2021 (cit. on p. 162).

- [289] Ahmed, H. O., Alfaqheri, T., Sadka, A. H., “Digital image inpainting techniques for cultural heritage preservation and restoration”, in *Data Analytics for Cultural Heritage*, Springer, 2021, pp. 91–122 (cit. on p. 162).
- [290] Kolomenkin, M., Leifman, G., Shimshoni, I., Tal, A., “Reconstruction of relief objects from line drawings”, in *CVPR 2011*, 2011, pp. 993–1000. DOI: [10.1109/CVPR.2011.5995643](https://doi.org/10.1109/CVPR.2011.5995643) (cit. on p. 163).
- [291] Setty, S., Mudenagudi, U., “Region of interest-based 3d inpainting of cultural heritage artifacts”, *Journal on Computing and Cultural Heritage (JOCCH)*, vol. 11, no. 2, pp. 1–21, 2018 (cit. on p. 164).
- [292] Barreiro Díaz, A., Munoz-Pandiella, I., Bosch, C., Andujar, C., “An Annotation Tool for Digital Restoration of Wall Paintings”, in *Eurographics Workshop on Graphics and Cultural Heritage*, The Eurographics Association, 2022. DOI: [10.2312/gch.20221220](https://doi.org/10.2312/gch.20221220) (cit. on p. 165).
- [293] Limoncelli, M., “Virtual restoration: Paintings and mosaics”, *Virtual restoration*, pp. 1–195, 2017 (cit. on p. 165).
- [294] Jam, J., Kendrick, C., Walker, K., Drouard, V., Hsu, J. G.-S., Yap, M. H., “A comprehensive review of past and present image inpainting methods”, *Computer vision and image understanding*, vol. 203, p. 103 147, 2021 (cit. on p. 166).
- [295] Gagniuc, P. A., *Markov chains: from theory to implementation and experimentation*. John Wiley & Sons, 2017 (cit. on p. 167).
- [296] Boiman, O., Shechtman, E., Irani, M., “In defense of nearest-neighbor based image classification”, in *2008 IEEE conference on computer vision and pattern recognition*, IEEE, 2008, pp. 1–8 (cit. on p. 168).
- [297] Chatfield, K., Philbin, J., Zisserman, A., “Efficient retrieval of deformable shape classes using local self-similarities”, in *2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops*, IEEE, 2009, pp. 264–271 (cit. on p. 168).
- [298] Shechtman, E., Irani, M., “Matching local self-similarities across images and videos”, in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2007, pp. 1–8 (cit. on p. 168).
- [299] Vedaldi, A., Gulshan, V., Varma, M., Zisserman, A., “Multiple kernels for object detection”, in *2009 IEEE 12th international conference on computer vision*, IEEE, 2009, pp. 606–613 (cit. on p. 168).
- [300] Schwarz, G., “Estimating the dimension of a model”, *The annals of statistics*, vol. 6, no. 2, pp. 461–464, 1978 (cit. on p. 168).
- [301] Brown, L. G., “A survey of image registration techniques”, *ACM computing surveys (CSUR)*, vol. 24, no. 4, pp. 325–376, 1992 (cit. on p. 168).

- [302] Student, “The probable error of a mean”, *Biometrika*, pp. 1–25, 1908 (cit. on p. [169](#)).
- [303] Kataoka, S., Yasuda, M., “Bayesian image denoising with multiple noisy images”, *The Review of Socionetwork Strategies*, vol. 13, no. 2, pp. 267–280, 2019 (cit. on p. [170](#)).
- [304] Rolland, J. P., Vo, V., Bloss, B., Abbey, C. K., “Fast algorithms for histogram matching: Application to texture synthesis”, *Journal of Electronic Imaging*, vol. 9, no. 1, pp. 39–45, 2000 (cit. on p. [171](#)).
- [305] Consortium, I. C., *Specification icc. 1: 2004-10 (profile version 4.2. 0.0) image technology colour management. architecture, profile format, and data structure*, 2004 (cit. on p. [171](#)).
- [306] Burt, P. J., Adelson, E. H., “A multiresolution spline with application to image mosaics”, *ACM Transactions on Graphics (TOG)*, vol. 2, no. 4, pp. 217–236, 1983 (cit. on pp. [171](#), [177](#)).
- [307] Borgefors, G., “Distance transformations in digital images”, *Computer vision, graphics, and image processing*, vol. 34, no. 3, pp. 344–371, 1986 (cit. on p. [173](#)).
- [308] Cornelis, B., Ružić, T., Gezels, E., Doms, A., Pižurica, A., Platiša, L., Cornelis, J., Martens, M., De Mey, M., Daubechies, I., “Crack detection and inpainting for virtual restoration of paintings: The case of the ghent altarpiece”, *Signal Processing*, vol. 93, no. 3, pp. 605–619, 2013 (cit. on p. [176](#)).
- [309] Daly, P., Evans, T. L., *Digital archaeology: bridging method and theory*. Routledge, 2004 (cit. on p. [183](#)).
- [310] Bodla, N., Singh, B., Chellappa, R., Davis, L., “Improving object detection with one line of code. arxiv 2017”, *arXiv preprint arXiv:1704.04503*, (cit. on p. [187](#)).