

Computational Analysis of Petroglyphs

DISSERTATION

submitted in partial fulfillment of the requirements for the degree of

Doktor der Technischen Wissenschaften

by

Markus Seidl

Registration Number 9255480

to the Faculty of Informatics

at the TU Wien

Advisor: Univ. Prof. Christian Breiteneder

The dissertation has been reviewed by:

Christian Breiteneder

Andreas Uhl

Vienna, 18th August, 2016

Markus Seidl

Computational Analysis of Petroglyphs

DISSERTATION

zur Erlangung des akademischen Grades

Doktor der Technischen Wissenschaften

eingereicht von

Markus Seidl

Matrikelnummer 9255480

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Univ. Prof. Christian Breiteneder

Diese Dissertation haben begutachtet:

Christian Breiteneder

Andreas Uhl

Wien, 18. August 2016

Markus Seidl

Declaration of Authorship

Markus Seidl
1020 Wien, Volkertplatz 4/20

I hereby declare that I have written this Doctoral Thesis independently, that I have completely specified the utilized sources and resources and that I have definitely marked all parts of the work - including tables, maps and figures - which belong to other works or to the internet, literally or extracted, by referencing the source as borrowed.

Vienna, 18th August, 2016

Markus Seidl

Erklärung zur Verfassung der Arbeit

Markus Seidl
1020 Wien, Volkertplatz 4/20

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 18. August 2016

Markus Seidl

Acknowledgements

First of all I would like to thank my wife Barbara and my son Lenz, both of whom supported me throughout my studies. Next, I want to thank my parents, my sister Simone and my brother Raimund for their support.

I want to thank my supervisor Professor Christian Breiteneder and my second supervisor Professor Axel Pinz for their continuous support, patience, motivation, insightful comments and encouragement.

I want to express my gratitude to the archaeologists Dr Frederik Baker and Dr Christopher Chippindale who let me participate in Cambridge University's Prehistoric Picture Project and introduced me to Valcamonica's cultural treasures in a rainy week in autumn 2010.

My sincere thanks go to all collaborators in the 3D-Pitoti project. First of all thanks to our fabulous coordinators Professor Sue Cobb, Dr Mirabelle d'Cruz and Sally Shalloe from the University of Nottingham. Tragically, Sally passed away unexpectedly and could not see the great success of the 3D-Pitoti project she had been working on. Sally, the memories of you are with me.

Special thanks go to Dr Alberto Marretta and Giovanna Bellandi from Archeocamuni, Tiziana Cittadini from the Centro Camuno di Studi Preistorici as well as Martin Schaich and Dr Oliver Reuß from ArcTron 3D for providing materials for this thesis. Furthermore, I want to thank Dr Christian Reinbacher and Georg Poier from Graz University of Technology as well as Craig Alexander from Cambridge University for the fruitful collaboration and joint publications.

In the 3D-Pitoti project I had the privilege to have Dr Matthias Zeppelzauer and Ewald Wieser in my team at the St. Pölten University of Applied Sciences. Our weekly discussions fostered the progress of the project and of my thesis. It was a pleasure to supervise Ewald's master thesis that contributed to the results presented in Chapter 8.

The research reported in this thesis has partially been carried out in the project 3D-Pitoti (<http://www.3d-pitoti.eu/>) which was funded from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 600545; 2013-2016.

Abstract

Numerous petroglyphs have been pecked, scratched and carved into rock surfaces in the northern Italian valley Valcamonica. The classic documentation work carried out by archaeologists is a massively time-consuming process. The rising availability of digital images and 3D scans of petroglyphs facilitates digital workflows which can improve the documentation process.

In this thesis, we aim at supporting the classic documentation pipeline for petroglyphs. The first step of the pipeline is the determination of the boundaries and spatial locations of petroglyphs on a rock surface. This is usually done by time-consuming manual contact tracing. Then, the found figures are classified according to their shapes and pecking styles. The large number of petroglyphs (Valcamonica contains up to 300.000 figures) demands large efforts for manual classification. The investigation of pecking styles is often impossible based on the contact tracings and thus requires researchers to return to the rocks frequently.

Following the classic documentation pipeline for petroglyphs, we propose and evaluate novel methods. To determine the positions and shapes of petroglyphs on a rock panel we approach segmentation of 2D and 3D petroglyph images in pecked regions and natural rock surface. Furthermore, we use 3D scans to investigate the similarity of pecking styles, i.e. the shape, size, depth and spatial distribution of the peck marks a figure consists of. Finally, we develop a petroglyph shape descriptor which allows the classification of petroglyphs. Our tasks are challenging. The figures have been pecked over thousands of years. The rocks are subject to weathering and abrasion. Therefore, the visual and tactile appearance of the petroglyphs varies greatly. Figures have often been superimposed over existing figures. Consequently, many merged and partial figures exist.

Contrary to previous work by others, we show that the segmentation of 2D images of rock surfaces is feasible. The employment of illumination-independent high-resolution 3D data of the surfaces' micro-topographies clearly improves results. We facilitate the investigation of pecking styles by modeling their similarity with 3D surface descriptors. The shape classification of a dataset containing more than thousand petroglyphs yields very good results with a combination of skeleton-, boundary-, and region-based shape descriptors. Our results can be useful for rock art researchers. Furthermore, we suggest how they can be applied in other domains.

Kurzfassung

Im norditalienischen Tal Valcamonica existieren zahlreiche Felsbilder, die in freiliegende Steinoberflächen gemeißelt worden sind. Die archäologische Dokumentation der Felsbilder (Petroglyphen) ist sehr zeitaufwändig. In den letzten Jahren wurden verstärkt Digitalphotographie und 3D Aufnahmetechnik zur Dokumentation der Petroglyphen eingesetzt. Die dabei entstehenden Bilddaten ermöglichen digitale Abläufe, welche wiederum den archäologischen Dokumentationsprozess verbessern können.

Die vorliegende Dissertation hat zum Ziel, Grundlagen für einen digitalen Dokumentationsprozess von Petroglyphen zu schaffen. Der erste Schritt der Dokumentation ist die Erfassung der Positionen der einzelnen Petroglyphen auf den Steinoberflächen. Dafür werden normalerweise die Petroglyphen mit einem Filzstift auf eine auf den Stein gelegte transparente Folie übertragen. In weiterer Folge werden die so aufgezeichneten Figuren nach ihren Formen klassifiziert. Nachdem alleine im Valcamonica bis zu 300.000 Petroglyphen existieren, ist die Klassifikation ein sehr arbeitsaufwändiger Prozess. Neben der Form ist die Machart der Gravuren (Form, Größe, Tiefe und Verteilung der einzelnen Meißelzeichen) ein wichtiges Merkmal der Petroglyphen. Der Gravurstil kann oft anhand der Folien alleine nicht beurteilt werden und erfordert so erneute Reisen zu den Steinoberflächen.

Wir entwickeln und evaluieren automatisierte Methoden für den Dokumentationsprozess von gemeißelten Felsbildern. Um das zeitaufwendige Durchpausen zu ersetzen, setzen wir neue Methoden zur Segmentierung von Petroglyphen aus Digitalbildern und 3D Aufnahmen ein. Wir verwenden die 3D Aufnahmen auch dazu, die Ähnlichkeit von verschiedenen Gravurstilen basierend auf 3D Deskriptoren der Oberfläche zu messen. Darüber hinaus entwickeln wir einen Formdeskriptor für Petroglyphen, der für die Klassifikation eingesetzt werden kann. Die Aufgaben sind sehr herausfordernd, da die Felsbilder über mehrere Jahrtausende hinweg in den Fels geschlagen wurden und die Felsoberflächen verschieden stark verwittert sind. Viele Petroglyphen sind außerdem über bereits bestehende graviert worden, daher gibt es viele Figuren, die nur in Teilen existieren.

Im Gegensatz zu früheren Arbeiten von anderen können wir zeigen, dass die automatische Segmentierung von Felsbildern basierend auf Digitalphotographien möglich ist. Die Verwendung der 3D Aufnahmen verbessert die Segmentierungsqualität, da die 3D Aufnahmen die Oberflächentextur beleuchtungsunabhängig erfassen können. Der Einsatz von 3D Oberflächendeskriptoren ermöglicht die automatische Unterscheidung

verschiedener Gravurstile. Wir evaluieren unseren Formdeskriptor auf einem Datensatz mit mehr als tausend Petroglyphen und erreichen sehr gute Klassifikationsergebnisse. Der Nutzen unserer Ergebnisse für Archäologen hat sich im Laufe der Evaluierungen gezeigt. Wir beschreiben darüber hinaus, wie unsere Ergebnisse in anderen Anwendungsbereichen sinnvoll eingesetzt werden können.

Contents

1	Introduction	1
1.1	Contribution	4
1.2	Organisation	6
2	Material	9
2.1	Rock Art	10
2.2	Rock Art Documentation	13
2.2.1	Documentation methods for Petroglyphs	13
2.3	Valcamonica’s Little Puppets - The Pitoti	15
2.3.1	(Re-)Discovery	16
2.3.2	Chronology	17
2.4	2D images of Petroglyphs	20
2.5	3D scans of Petroglyphs	20
2.6	Ground Truth	24
2.6.1	Segmentation Annotation	26
2.6.2	Shape Annotation	27
2.6.3	Shape Annotation Tool	29
2.6.3.1	Main Features	29
2.6.3.2	Handling and Annotation Workflow	31
2.6.3.3	Statistical Reporting	34
2.7	Conclusion	35
3	Petroglyph-related Work	41
3.1	3D Scanning of Petroglyphs	42
3.2	Segmentation of Petroglyph Images	44
3.3	Pecking Style Similarity	45
3.4	Shape Classification of Petroglyphs	45
3.5	Discussion	46
4	Segmentation of Petroglyph Images	49
4.1	Related Work	51
4.2	Approach	51
4.2.1	Pixel Classification	51

CONTENTS

4.2.2	Fusion Strategies	52
4.3	Experiments	53
4.3.1	Material	53
4.3.2	Visual Features	55
4.3.3	Training and Classification	56
4.4	Results	58
4.4.1	Single Features	58
4.4.2	Early Fusion	58
4.4.3	Late Fusion	60
4.4.4	Late Fusion with Thresholded Majority Voting	60
4.4.5	Interactive Late Fusion	63
4.5	Conclusion	64
5	Segmentation of Petroglyph 3D Point Clouds	67
5.1	3D Descriptor-based Point Cloud Segmentation	68
5.1.1	Related Work	68
5.1.2	Approach	68
5.1.3	Results	71
5.2	Image-space-based Point Cloud Segmentation	71
5.2.1	Related work	73
5.2.2	Descriptors	74
5.2.3	Approach	74
5.2.4	Experimental setup	75
5.2.5	Results	77
5.2.5.1	Small dataset	77
5.2.5.2	Full dataset	80
5.3	Conclusion	81
6	Similarity of Pecking Styles	85
6.1	Related Work	86
6.2	Approach	89
6.2.1	Dense sampling	89
6.2.2	Topology-based sampling	89
6.3	Results	92
6.3.1	Dense sampling	92
6.3.2	Topology-based sampling	95
6.4	Conclusion	99
7	Description of Petroglyph Shapes	103
7.1	Related Work	103
7.1.1	Region-based Descriptors	104
7.1.2	Contour-based Descriptors	105
7.1.3	Skeleton-based Descriptors	105
7.1.4	Discussion of Shape Descriptors	107

7.1.5	Graph Matching	107
7.2	Proposed Approach	108
7.2.1	Benchmark Dataset	108
7.2.2	Overview of our Approach	108
7.2.3	Descriptors	111
7.2.4	Descriptor Fusion and Classification	111
7.3	Experimental Results	113
7.3.1	Graph Matching	113
7.3.2	Comparison with Baseline Descriptors and Descriptor Fusion . .	114
7.4	Conclusion	116
8	Skeletonisation of Petroglyph Shapes	119
8.1	Related work	121
8.1.1	Point-based pruning approaches	122
8.1.2	Branch-based pruning approaches	122
8.1.3	Comparison of algorithms	123
8.2	Approach	124
8.2.1	Investigated material	126
8.2.2	Adaptive shape pre-processing	126
8.2.3	Improvements of existing skeletonisation methods	128
8.3	Experimental setup	131
8.3.1	Selection of parameters	131
8.3.2	Evaluation	132
8.4	Results	132
8.5	Conclusion	137
9	Automated Classification of Petroglyph Shapes	139
9.1	Approach and Related Work	140
9.2	Experiments and Results	141
9.2.1	Descriptor Combination Previously Proposed	141
9.2.1.1	Modified Sansoni Typology	141
9.2.1.2	3D-Pitoti Project Typology	145
9.2.2	Details on Single Shape Descriptors and CNNs	149
9.3	Use Scenario	149
9.4	Conclusion	152
10	Conclusion and Future Work	155
10.1	Conclusion	156
10.2	Future Work	157
	List of Figures	159
	List of Tables	167

CONTENTS

References	169
Appendices	185
A 3D dataset	187
B Peer-reviewed Publications	211

Introduction



Figure 1.1: A petroglyph depicting a warrior in Valcamonica, Italy.

PETROGLYPHS have been pecked, scratched and carved into rock panels all over the world since thousands or even ten thousands of years. They are among the oldest cultural traces of mankind and allow us to have a glimpse as far as to the early stages of human civilisation. Although subject to weathering and abrasion of the rock surfaces they are more stable and longer lasting than painted rock art (pictographs). The petroglyphs are not made of material but exist because of material that has been pecked or scratched away from a rock surface. One consequence is that they cannot be dated objectively by Carbon dating. The second consequence is that, as they are existing at the surface of a rock, they are flat (or slightly curved with the curvature of

1. INTRODUCTION

the rock). This flatness led to a graphical style that combines different perspectives. If we look at the warrior in Figure 1.1, we notice that some body parts have been rotated in order to be depicted and recognised. The feet are pointing to the right while the whole figure with equally long left and right arms and legs suggests a depiction in frontal perspective.

Petroglyphs stem from the past and are evidence of mankind. The archaeologists researching different aspects of petroglyphs are *rock art researchers*. To be able to infer knowledge from evidence stemming from the past, the evidence must be documented, sorted, archived and preserved for future researchers. The classic documentation method for petroglyphs is time-consuming manual contact tracing on transparent sheets (see Figure 1.2). The result of such a tracing is depicted in Figure 1.3. Digital photography and image stitching methods already moved forward the documentation of rock art. With the rising availability of affordable 3D scanning techniques, more and more rock surfaces containing petroglyphs will be scanned and subsequently be available digitally as highly detailed 3D scans. The availability of 3D scans of rock art is not necessarily an immediate advantage for the rock art researcher and the interested public. In order to exploit the potential, the 3D data must be acquired with geo-referencing, must be stored safely, must include tools for navigation and allow a viewing experience as good as the viewing experience directly on site. Moreover, the data should also improve the work of rock art researchers and allow new insights which have not been possible before. The main aim of this thesis is to investigate whether digital 2D and 3D depictions of petroglyphs and their *(semi-)automated semantic enrichment* can enable *new methods for rock art documentation and analysis*.

With the vision of large archives full of rock art images and 3D scans we aim at robust automated methods to

- (a) determine the exact shapes and spatial positions of petroglyphs in images or 3D scans of full rock panels (i.e. segmentation of the image or 3D scan in pecked and non-pecked regions);
- (b) classify the petroglyphs regarding their shapes and pecking styles and
- (c) retrieve similar petroglyphs from different archives of petroglyph images and 3D scans.

We have access to high-resolution photographs and high-resolution 3D scans of rock surfaces. Furthermore, we have many classical petroglyph tracings of full rock panels. The material stems from the UNESCO world heritage site *Valcamonica* in northern Italy.

The motivation to bring computer vision methods to the domain of rock art research is twofold. First, to address human solvable problems that occur in high number and are therefore repetitive. Second, to solve problems, that are hard to solve for humans, e.g. because of visibility and/or scale. We clearly have the first type of problems in our case: Our test-bed Valcamonica contains 300.000 petroglyphs. The documentation of the shape and the spatial location of all figures is a massively time-consuming process that is currently far from being completed. The subsequent step of classification of



Figure 1.2: Manual contact tracing in Valcamonica, Italy. ©Alberto Marretta, used with permission.

1. INTRODUCTION

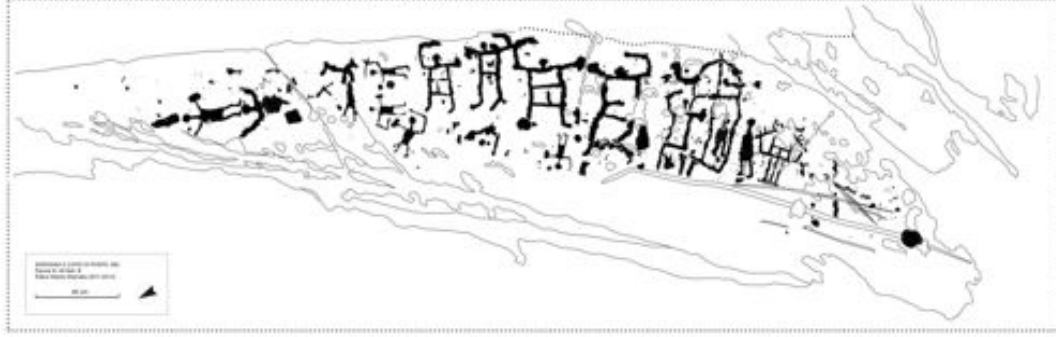


Figure 1.3: Manual contact tracing from Seradina Rock II in Valcamonica, Italy.²

the figures into archaeological typologies is open for most of the rocks. Hence, the automated segmentation of rock surfaces in pecked and natural areas as well as the automated classification of the segmented figures has the potential to clearly enhance the documentation and classification work of rock art researchers. The second type of problem occurs in connection with the pecking styles, i.e. the shape, depth and spatial distribution of the single peck marks a petroglyph consists of. The high-resolution 3D scans with a precision of 0.1mm can enable a very close look on single peck marks and may allow an automated similarity estimation of differently pecked regions. The knowledge about the similarity of pecking styles in different regions can help in relative chronology. In the case of two merged figures usually one figure is superimposed over the other. If the pecking styles of the two figures can be distinguished, the rock art researcher can infer that the figure superimposed over the other has been pecked later.

The research for this thesis was partly carried out in the 3D-Pitoti project³. The project was realised in an international consortium with seven partners. It was dealing with a completely 3D-based digital rock art pipeline from scanning over intelligent processing to interactive visualisation (see Figure 1.4). The test-bed for the tools developed in the project is the UNESCO world heritage site Valcamonica.

1.1 Contribution

This thesis aims to lay a fundament for a digital toolset to analyse petroglyphs based on high resolution 3D scans. However, our contributions have impact outside the rock art domain, too. In the following paragraphs we line out our main contributions and their impact beyond the scope of the thesis. The scientific publications based on materials

²©Alberto Marretta, used with permission.

³This project was funded from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no 600545; 2013-2016. For more information about the project please visit <http://3d-pitoti.eu>.

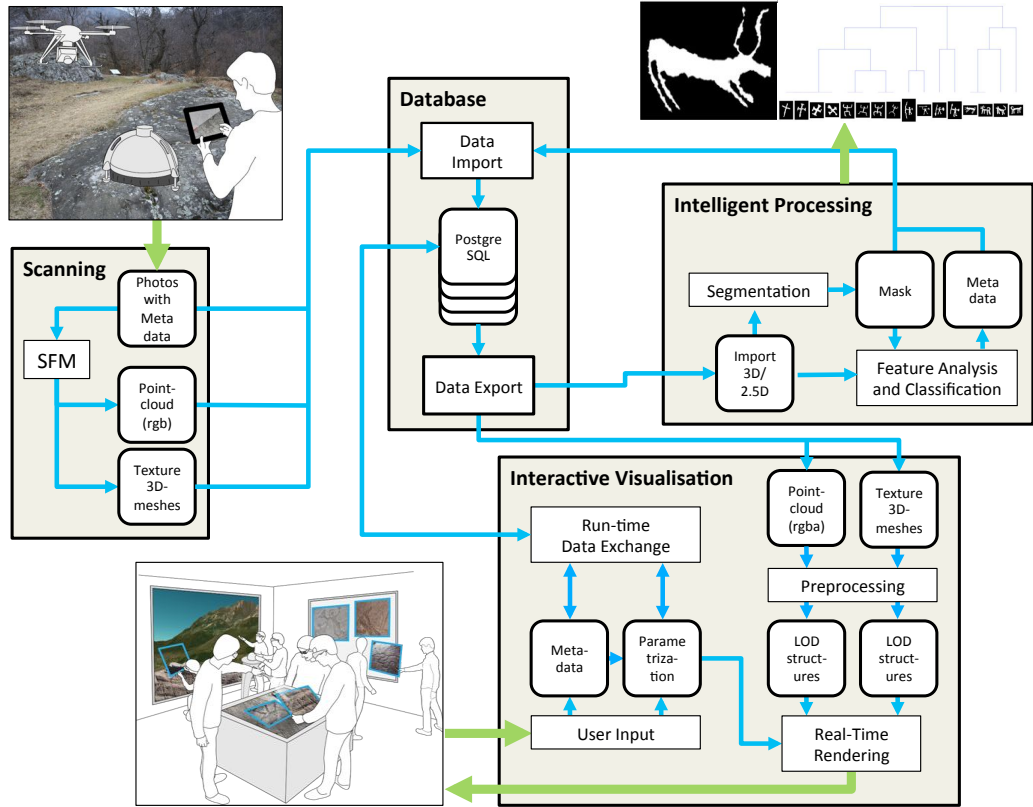


Figure 1.4: Simplified system workflow of the complete 3D-Pitoti pipeline. The scanning was carried out as multi-scale approach based on the 3D-reconstruction method Structure-from-Motion (SFM). The massive amount of multi-scale 3D-data required Level-of-Detail (LOD) algorithms for interactive visualisation. The contributions of this thesis (see Section 1.1) are subsumable under “Intelligent Processing”. We thank Alex Kulik for the preparation of the first version of this figure.

1. INTRODUCTION

of this thesis are listed in Appendix B. We can provide all data used in this thesis upon request.

Segmentation: For the time-consuming first step of rock art documentation - tracing - we propose *digital tracing*. For a task that has even been considered infeasible by other researchers ([ZWKL11]) we propose an approach that segments a point cloud (or image) in foreground (pecked rock surface) and background (natural rock surface). First, we demonstrate that segmentation of petroglyphs can be performed with high quality digital photos as input. The main disadvantage of using photos for segmentation is the dependency on the lighting situation during acquisition. To be independent of the lighting situation we use high resolution 3D scans of the rock surfaces in the next step. We demonstrate that segmentation based on 3D scans outperforms segmentation of petroglyph photos. But, the dense extraction of 3D point cloud descriptors is computationally demanding. Thus, in a next step we move the problem to image space by using the depth map of the rock surface as input. We show, that the depth map-based method yields the best results. Our approach and its evaluation could have impact outside the rock art domain as the availability of high-resolution 3D surface scans will probably increase strongly and texture classification based solely on the 2D acquired *visual appearance* of a surface might be outperformed or even replaced by texture classification based on 3D data, i.e. the lighting independent *tactile appearance* of a surface.

Classification: Subsequent to the tracing of the petroglyphs on the rock surface the single petroglyph figures are classified into pre-defined typologies. We provide a petroglyph shape similarity measure based on the combination of a) a novel descriptor based on the skeletal graph of a petroglyph, b) shape descriptors used by other researchers for petroglyph shape classification and c) state-of-the-art image features based on convolutional neural networks. We evaluate our approach on a large shape database containing more than thousand petroglyphs. Our approach outperforms all other approaches. We provide the evaluation data to the community for further investigations. Our approach and evaluations have impact outside the rock art domain as we on one hand investigate the question “*Who needs shape descriptors in the age of convolutional neural networks where features are simply learned?*” and on the other hand existing datasets for shape classification are highly artificial compared to our shape data created by humans.

3D-Data: Most of the tasks described in this thesis are based on high-resolution 3D data. But, the reader will recognise that we transferred tasks to 2D image space. Thus, this thesis implicitly investigates the general question “*How to handle the computational analysis of large amounts of high-resolution 3D data?*”.

1.2 Organisation

The organisation of the thesis is as follows. Subsequent to the introduction, we extensively describe our material in Chapter 2. We define rock art in Section 2.1 and introduce classic rock art documentation in Section 2.2. In Sections 2.3 to 2.6 we describe our material and the ground truth acquisition process. The related work is split

in related work dealing with petroglyphs which is presented in Chapter 3 and technical related work which is - with the aim of better readability - summarized in each technical Chapter. Chapters 4, 5 and 6 deal with the characteristics of the *micro-topography of the rock surface*. Chapter 4 presents an approach for the segmentation of rock surfaces in pecked and non-pecked regions based on digital photos. Chapter 5 deals with the same problem but uses 3D point clouds as input. Chapter 6 describes investigations on the similarity of different pecking styles. Chapters 7, 8 and 9 deal with the *shape of petroglyphs*. Chapter 7 presents an approach to describe and match petroglyph shapes based on their skeletal graph. The quality of the skeletal graph is depending on the quality of the skeletonisation. Hence, in Chapter 8 we investigate the skeletonisation and other necessary pre-processing steps for shape description of petroglyphs. In Chapter 9 we evaluate the findings of the previous chapters on a large petroglyph shape dataset and compare and combine our descriptor with other shape descriptors and an additional state-of-the-art image matching method. Finally, in Chapter 10 we reflect, draw conclusions and line out future work.

1. INTRODUCTION

Material



Figure 2.1: A petroglyph scene in Valcamonica, Italy.

IN this chapter we describe the origin of our material and our ground truth data. Section 2.1 contains the definition of rock art, its world wide occurrences and some prominent examples. Section 2.2 describes documentation issues of rock art focusing on the documentation of petroglyphs. Section 2.3 presents the rock art from Valcamonica. Sections 2.4 to 2.6 describe our ground truth data and the annotation process.

2. MATERIAL

2.1 Rock Art

Withley defines rock art as follows [Whi05]:

Rock art is landscape art. It consists of pictures, motifs, and designs placed on natural surfaces, such as cliff and boulder faces, cave walls and ceilings, and the ground surface. Rock art is also sometimes referred to as cave art or parietal (wall) art. Regardless of designation, the defining characteristic of rock art is its placement on natural rock surfaces.

Rock art is a global phenomenon. It can be found all over the world. Many rock art sites are part of the UNESCO world heritage. The UNESCO ICOMOS documentation centre lists rock art sites in Africa, the Arabian States, Asia and the Pacific, Europe as well as Latin America and the Caribbean [Sil09]. In total, 31 rock art sites are a part of the UNESCO world heritage. Rock art belongs to the oldest traces of human culture. Pike et al. dated rock art from 11 caves in Spain based on uranium-series disequilibrium dates of calcite deposits overlying or underlying art [PHGD⁺12]. The oldest piece of rock art they dated is at least 40.800 years old. Also outside of Europe, old rock art has been found and dated. Aubert et al. dated cave art in Sulawesi, Indonesia using uranium-series dating of coralloid speleothems [ABR⁺14]. They found a hand stencil with a minimum age of 39.900 years.

The Lascaux cave in France contains a well-known example of European rock art. The paintings in the cave are dated around 15.000 BC. Hence, they are at least 17.000 years old. They show animals that the Paleolithic humans had in their surroundings, both animals that were hunted and eaten, such as deer and bison as well as animals that were feared, such as bears and wolves. The animals are shown in a twisted perspective, a combination of front view and profile which allows the depiction of the imposant profile of a bison with both horns (see Figure 2.2). The images have been painted with pigments of minerals readily available. However, no brushes have been found. Most probably the black outlines have been painted with moss or hair or even the raw pigments. The surfaces could have been covered with pigments blown to the surface, as hollowed-out bones internally covered with pigments have been found in the caves [Ted00]. Generally, rock art can be subsumed under three basic types [Whi05]:

- *Petroglyphs* are pecked or scratched into rock surfaces (see Figures 2.1 and 2.3). The peckings are sometimes called engravings or carvings. They are made with fist-sized stones that are used as hammerstones. A stone of a certain type (e.g. basalt) can be used as hammerstone on the same type of rock. Another way of making petroglyphs is to use a fist-sized stone or a metal tool as chisel on which one hammers either with another stone or with a piece of wood. The rarely occurring scratched petroglyphs are made with a lithic flake or blade.
- *Pictographs* are drawn or painted onto rock surfaces (see Figure 2.2). They are made of common materials in the whole world. The materials are minerals and other natural substances. The often used red is made of ground ocher, black is



Figure 2.2: Cave painting in the Lascaux cave in France showing superimpositions of two bison bulls, horses, deer and other animals.²

mostly made of charcoal. The materials are either used dry (e.g. charcoal to draw on rock) or ground and mixed with water and then put wet on the rock surface.

- *Earth figures* are large designs or motifs that are on the ground surface. They are divided in two categories depending on the type of creation. Motifs like the Nazca lines in Peru (see Figure 2.4) that are made by removing the surface of desert ground and creating a negative image are called *Intaglios*. Forms that are created as positive images on the ground by moving rocks are called *geoglyphs*.

The main concern regarding rock art is its ephemeral nature. Unless it is in a cave, it is exposed to nature and thus subject to weathering. Depending on the rock they are pecked into, Petroglyphs are relatively insensitive to the environmental conditions. However, pictographs (painted rock art) can be washed away within one rainy season. Hence, painted rock art is mostly found in caves or in dry regions. Thus, a central point, the step zero of rock art research, is documentation of rock art.

²©Prof Saxx, CC-BY-SA-3.0, source: https://en.wikipedia.org/wiki/Lascaux#/media/File:Lascaux_painting.jpg.

2. MATERIAL



Figure 2.3: A petroglyph depicting a Bubalus in the Sahara in southern Algeria. ©Linus Wolf, CC-BY-SA-3.0.



Figure 2.4: Intaglio showing a hummingbird in Nazca, Peru. The figure is almost 100 meters long.³

2.2 Rock Art Documentation

Basically, rock art is documented with pictures, e.g. with (multispectral) images of pictographs or aerial images of geoglyphs. Petroglyphs have a third dimension which is of interest. How are the pecks or scratches formed? How deep are they? Which tool could have been used? Different techniques to document petroglyphs have been developed to answer these questions.

2.2.1 Documentation methods for Petroglyphs

Rock art researchers spend a considerable amount of time to document the rock engravings. The first technique developed for the documentation of petroglyphs was *frottage* (engl. rubbing). A sheet of paper is placed on the petroglyphs and a dyeing material (e.g. grass, charcoal or graphite) is rubbed onto the surface of the paper. The edges on the rock surface create different intensities of color. See Figure 2.6 for the process of rubbing and Figure 2.5 for a resulting image of a whole panel.

Another technique is *casting* with plaster which results in a negative mask of the rock surface and therefore delivers a very precise depiction of the three dimensional surface. But, it is obviously a very time-consuming technique. The most frequently used technique since the 1950ies is manual *tracing*. A transparent foil is put on the rock surface and each peck mark is traced with a marker pen. Tracing delivers - as long as the tracing person is concentrated - a very precise depiction of the structure of the rock art and the characteristic of the pecking - the pecking style. Figures 2.7 and 2.8 show

³©Bjarte Sorenson, CC-BY-SA-3.0, source: http://commons.wikimedia.org/wiki/File:Nazca_colibri.jpg.

2. MATERIAL



Figure 2.5: Result of a rubbing documentation. The rubbing shows a rock panel in Tanum, Sweden. ©Gerhard Milstreu / Tanums Hällristningsmuseum, used with permission.



Figure 2.6: The process of rubbing on a rock panel in Tanum, Sweden.⁴

the tracing process while Figure 1.3 on page 4 shows the result of such a tracing.

Besides photography, which has always been used to document petroglyphs and digital photography, which extended the possibilities about 20 years ago, 3D scanning of petroglyphs has been performed since around 10 years ago. The results that can be obtained with 3D scanning can be seen as a combination of classic photography with the documentation techniques mentioned above to record the pecks and scratches in their three-dimensionality. We describe the 3D scanning of petroglyphs in more detail in Chapter 3.

2.3 Valcamonica's Little Puppets - The Pitoti

More than 300.000 “prehistoric” figures have the immediacy and the historical reality of documents created by those who participated in the very events which led to the formation of European civilization, and which were left, not as chronicles for posterity, but as an organic function of their lives. In this way they act as a medium through which we can revive the past.

Emanuel Anati [AC94]

The *Valle Camonica* or, more frequently used, *Valcamonica* is an alpine valley in Lombardy (a northern Italian province) stretching north from the Lago d’Iseo and terminating at Passo Tonale. It is a UNESCO world heritage site with numerous rock panels containing more than 300.000 petroglyphs stemming from the last 8000 years. Most of the petroglyphs originate from prehistoric times. They have been out of focus outside the valley until the 1930ies, when first (re-)discoveries of pecked surfaces were made. In 1979 the Valcamonica rock art was included in the UNESCO world heritage list. In

⁴©Attie Logger / Tanums Hällristningsmuseum, used with permission.

2. MATERIAL



Figure 2.7: Tracing on a rock in Valcamonica in Italy.⁵

1994 Anati published the book “Valcamonica Rock Art - A New History of Europe” [AC94] on which the following sections are based unless otherwise noted.

2.3.1 (Re-)Discovery

The Romans conquered Valcamonica in the first century AD. Subsequently, most of the pre-Roman settlements and with them the petroglyphs were abandoned and forgotten. Over 2000 years, moss and earth grew over most of the rock panels. Only a few rock panels were still at least partly visible in the 19th century. One of these was known as *Preda die Pitoti*, the puppets’ rock.

In 1930, two scholars from the University of Florence and the University of Turin visited the puppets’ rock in Capo di Ponte separately. Both published on this one rock including photographs and descriptions of the engravings. Following these publications, in the 1930ies, several more rock panels with petroglyphs have been found and documented, among these inscriptions in an unknown language using Northern Etruscan characters. Franz Altheim, a Professor of history at the University of Berlin and supporter of Hitler’s racist theories came to the valley to find traces of the Aryan civilizations. He published a work “Vom Ursprung der Runen” (1933) which has a pref-

⁵©Alberto Marretta, used with permission.

ace by Heinrich Himmler. In WW II and the subsequent years, only little research was carried out in Valcamonica. In the 1950ies, Gualtiero Laeng and Emanuele Suss set out to search for rock panels with engravings. Suss made the first map of rock panels of a central area (the Naquane area) containing around 100 rock panels with engravings. This site was then almost completely documented - only six more rocks have been found in this area since then. In the rest of the valley more than 2500 rock panels have been found from 1956 to 1994. The “Anati Expedition” discovered 30.000 Pitoti from 1956 to 1964, when the Centro Camuni Studi Preistorici (CCSP) has been founded by Emmanuel Anati. The CCSP carried on the excavation and documentation of Valcamonica petroglyphs, which are called “Pitoti” (plural, singular is “Pitoto”) a local dialect word for “little puppet”.

2.3.2 Chronology

The variation of figures found in Valcamonica is high (see Figure 2.15 on page 28). Some figures try to depict reality as close as possible, some are highly stylized, while others are abstract. The work techniques used to make the petroglyphs vary, too. Most petroglyphs have been pecked rather than scratched or carved. The process of pecking was performed directly with a stone in the fist or indirectly with a chisel and a hammer. The chisel was made of stone or metal while another stone or a piece of wood was used as hammer. A basic problem of all attempts to rock art chronology is the fact, that no dating method like the investigation of the employed painting materials or C14 analysis can be used. Hence, each chronology has the risk of being questioned. Marretta [Mar11] describes the dating approach as a two step process. First, a relative chronology between two figures is established by the analysis of superimposition which exploits the fact that a figure pecked over another figure must have been produced later. Second, an absolute chronology is derived by the identification of similarities between the figures and corresponding real objects, e.g. weapons, plows, wagons, etc. that have been found and dated absolutely with stratigraphic methods and C14 analysis. Additionally, the consideration of extinct animals and the knowledge about the coming to existence of certain objects (e.g. the plow in the third millenium BC) can support the dating process.

In [AC94] Anati argues, that seven properties that are unique in Valcamonica compared to other rock art sites enabled the chronology.

1. The high number of petroglyphs that have been found in Valcamonica delivers sufficient data.
2. The high amount of superimposition, i.e. figures that have been pecked over each other, help to establish relative chronologies between the superimposed figures as it can be seen in most of the cases which figure was pecked over the other, i.e. pecked later.
3. The large number of depicted weapons and other objects with a shape distinct enough to be matched with real and hence C14 datable objects found in excavations enables absolute chronological cues.

2. MATERIAL



Figure 2.8: Tracing on a rock in Valcamonica in Italy. ©Alberto Marretta, used with permission.

2.3 Valcamonica's Little Puppets - The Pitoti

4. Variations in style allow to group figures.
5. Excavated human settlements containing rocks with engravings allow to date the rocks through stratigraphy.
6. Inscripted Etruscan letters indicate time frames in which the letters must have been pecked.
7. Animal figures of extinct species give an absolute cue of the last possible date when the figure must have been pecked.

The chronological analysis based on the seven criteria led to the first chronology by Anati [Ana76]. The chronology consists of six principle periods. First, a Proto-Camunian period contains a hunter-gatherer society. The four subsequent periods cover the actual Camunian civilization, a food production lifestyle. The sixth period is the Post-Camunian period containing Roman and Post-Roman time.

2. MATERIAL

2.4 2D images of Petroglyphs

In the previous sections, we have presented general information about our material. In the remaining sections of this chapter we will focus on the data we use for our tasks *segmentation* of rock surfaces containing petroglyphs and *classification* of the resulting figures. The first step in documentation of petroglyphs is the recording of the spatial locations of the petroglyphs on a rock surface. The classic documentation methods (rubbing, casting and tracing, see Section 2.2) rely on labor intensive work directly on the rock surface. For our initial investigations on segmentation of petroglyph depictions in natural rock surface and pecked rock surface we worked with high quality digital photography. The pictures have been taken with a Canon EOS 5 MK II and different high quality original Canon lenses. Figures 2.9(a) and 2.10 show parts of a pecked surface. The pictures have sufficient resolution to show detailed depictions of the local surface (see enlarged crop in Figure 2.9(b)).

However, in practice there is one central obstacle of working with 2D photographs that strongly influences the quality of the results: The dependence on the lighting situation. Figure 2.9(a) shows an example with shadows from trees. Additionally to the problem of other objects casting shadow on the surface, the incoming light direction strongly influences how the three-dimensional micro-structure of a surface is depicted on a photograph. On an overcast day with ambient lighting the surface structure is less visible than on a sunny day in the morning or in the afternoon, when the sun rays are close to parallel to the surface. Near parallel illumination is beneficial for image-based 2D approaches as it generates stronger shadows and thus makes the surface geometry visible better. This problem can only be overcome by the use of a flash light. A flash device mounted on top of the camera is, however, not suitable as the light rays of the flash would be close to orthogonal to the surface and, hence, reduce visibility of the geometric micro-structure of the surface. The application of a mobile studio flash light improves the situation, as the flash can be located in an adequate angle to the surface. The problem that remains is the high power consumption of the flash which requires carrying large and heavy battery packs to the capture site. This is often not feasible in remote areas and on steep surfaces. Therefore, we captured our 2D material without a flash light.

2.5 3D scans of Petroglyphs

In the previous section we described our 2D photographic material and its main disadvantage - the lighting dependency during acquisition. An alternative way to capture the micro-topography of the surface is 3D scanning, which is frequently used for archaeological documentation.

We work with a dataset acquired by the german 3D-surveying specialist ARCTRON¹ for the 3D-Pitoti project. The dataset contains 26 surface reconstructions with a resolution of 0.1mm. The number of points per surface ranges from 810k to 10.7M. An

¹<http://www.arctron.de>



(a) Full image.



(b) Detail. We observe, that the peck marks can be seen well.

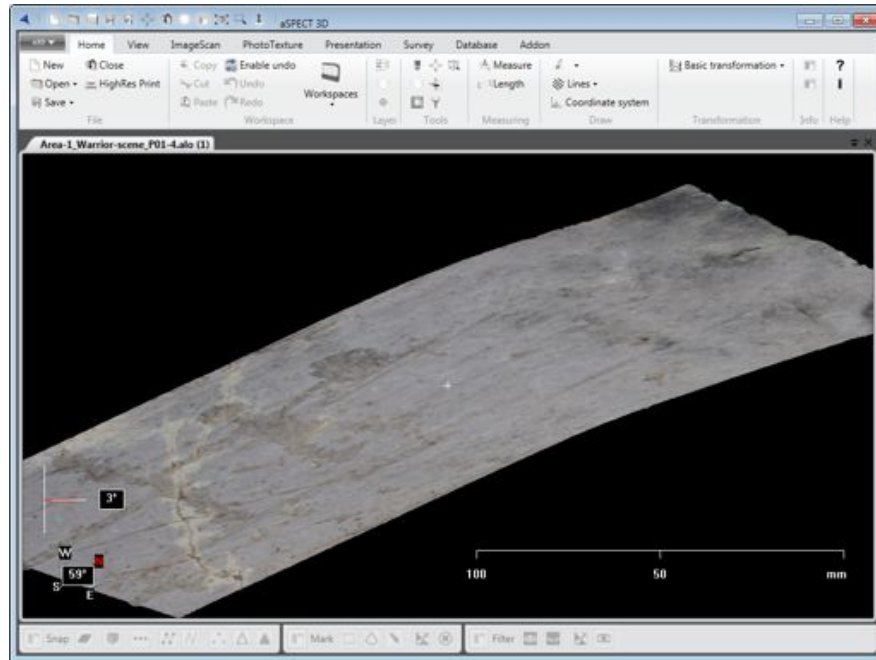
Figure 2.9: High quality 2D images as used in 2D segmentation. Shadows of trees surrounding the rock are visible.

2. MATERIAL

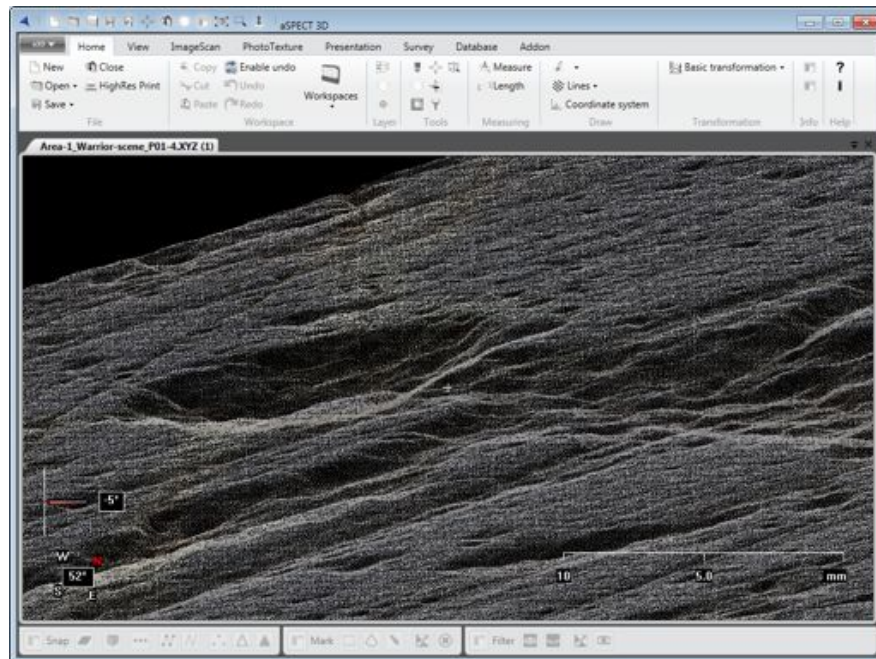


Figure 2.10: A high quality 2D image as used in 2D segmentation.

2.5 3D scans of Petroglyphs



(a) Point cloud.



(b) Detail.

Figure 2.11: A 3D input point cloud (a) and a close up of the head of the Pitoto (b). The point cloud has 4.4M points. The close up clearly shows the surface morphology of the pecked regions (e.g. in the head and torso of the Pitoto).

2. MATERIAL

example for a point cloud with 4.4M points together with a close-up zoom into the cloud is shown in Figure 2.11. See Appendix A for details on the 3D dataset.

From the input point clouds, we computed a number of derived representations suitable for different purposes. The derived representations include:

- Exports of the point clouds in the XYZRGB¹ format,
- Orthophotos of the point clouds with oblique lighting,
- Depth maps registered with the orthophotos,
- Parameters of the orthographic transform.

The XYZRGB format has been selected due to its simple and easy to process format. Each point is stored together with its color values in one row of a plain text file. The orthographic projection (together with the corresponding depth map) of the surface gives us an easy processable representation of the 3D surface. Since the surface scans do not contain self-occlusions or overhanging structures, there are no occluded areas in the orthographic projection. Thus, no information available in the 3D scan gets occluded in the projection. Consequently, the orthographic projection is an adequate representation of the 3D scan. For the orthographic projection, we take an input cloud and fit a median plane through the cloud. Once the median plane is found the projection direction is set to the normal direction of the median plane. The result of orthographic projection is a projection of the point cloud to a 2D image (orthophoto) that is parallel to the median plane. The process for depth maps is in general the same as for the generation of orthophotos. The only difference in computation is that instead of mapping the color of the point at the projected pixel position in the orthophoto, the distance of the point from the median plane is mapped to the pixel position. The resulting depth maps are stored as 32bit TIFF images to preserve the full depth precision provided by the point cloud.

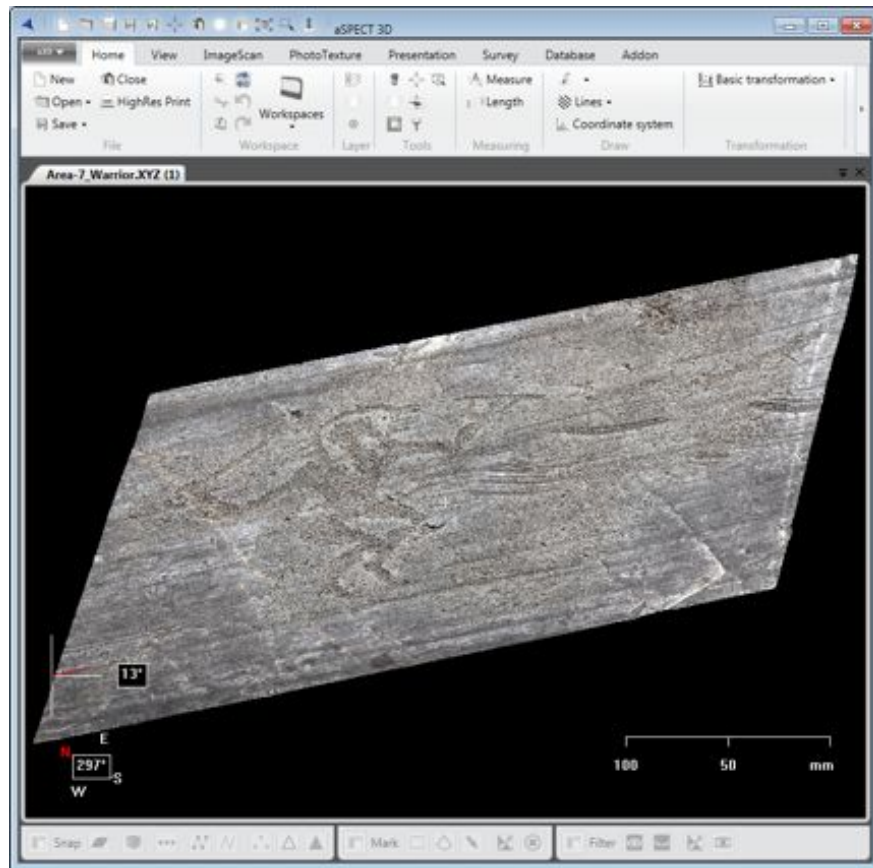
The point clouds have a resolution of 0.1mm. They are stored with metric coordinates. The orthographic projection is performed using the software Aspect3D from ARCTRON². The image resolution is set to 300dpi (equivalent to 0.08 pixels per mm), which has shown to be sufficient for the generation of accurate orthophotos and depth maps.

2.6 Ground Truth

In order to be able to quantitatively evaluate our methods we need to enrich our data with expert knowledge about it. This gold standard is called the *ground truth*. We compare the output of our algorithms with the ground truth to evaluate the algorithms' performance. For the methods that aim at segmentation in natural rock and pecked surface areas we need annotations of the surface containing information whether a specific pixel or 3D point is considered foreground or background (see Subsection 2.6.1).

¹The XYZRGB format stores a sextuple for each 3D point containing the spatial coordinates (x,y,z) and RGB color values (r,g,b).

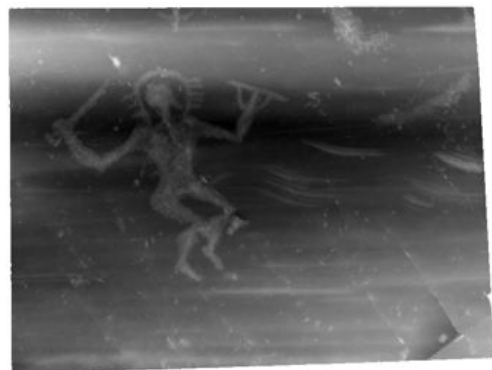
²<http://aspect.arctron.de/>



(a) Point cloud.



(b) Ortho photo.



(c) Ortho-depth photo.

Figure 2.12: An input point cloud (a) with its orthophoto (b) and depth map (c). Brighter values in the depth map indicate higher depth values.

2. MATERIAL



Figure 2.13: Segmentation ground truth annotation on the 2D image in Photoshop.

For the ground truth on the types of the petroglyphs we need a much larger number of examples than the few figures depicted in detail in the 26 3D scans. We use classic tracings of Valcamonica petroglyphs for this purpose and have a database containing more than 1000 expert-annotated figures in two different typologies (see Subsection 2.6.2). To establish this database we developed a user friendly web-based tool (see Subsection 2.6.3).

2.6.1 Segmentation Annotation

For the evaluation of the segmentation algorithms, we asked domain experts to create manual segmentations that provide a ground truth for the evaluation of the performance of the developed algorithms. All pixels in an orthophoto being part of a pecked region were annotated. The result is a binary mask which is true for pecked regions and false for natural rock surface.

The annotation was performed by Archaeology experts with image processing tools like Gimp and Adobe Photoshop. For the 2D images we used the photos in Photoshop (see Figure 2.13). Manual annotation of 3D surfaces requires special tools. As our surfaces do not contain self-occlusions, we used orthophotos of the 3D surfaces for the annotation. The process was the same as for the 2D images. For ambiguous areas of the orthophotos we additionally provided a 3D viewer for the annotating experts. All 26 orthophotos have been precisely annotated. Examples of 3D ground truth annotations are shown in Figure 2.14.

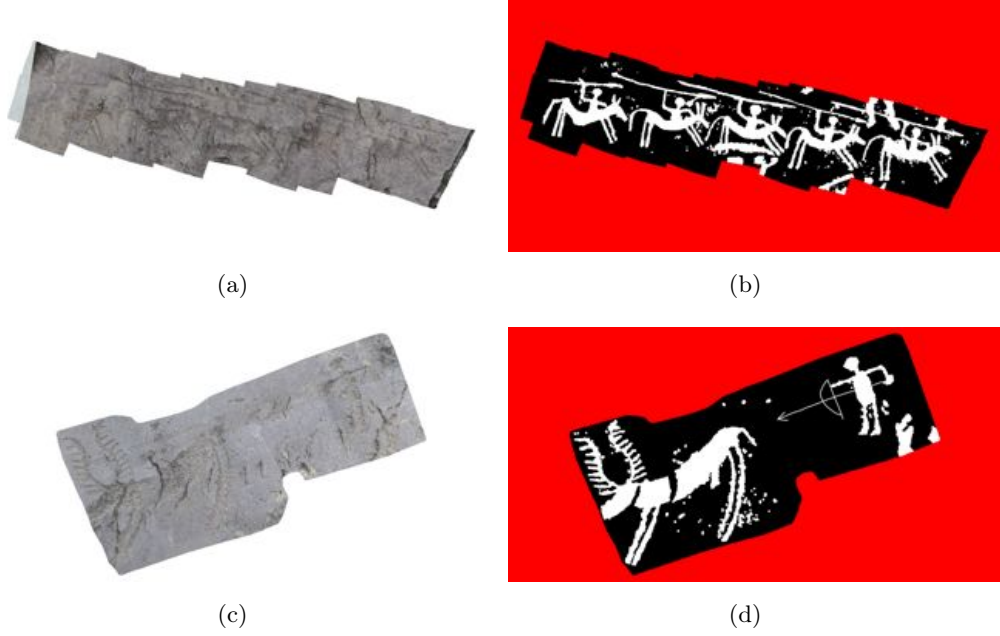


Figure 2.14: Orthophotos (a and c) and their corresponding ground truth images (b and d). White areas mark pecked regions in the orthophoto and black areas unpecked regions. The red area surrounding the scan is the area where no information is available.

2.6.2 Shape Annotation

The purpose of the shape annotation is the quantitative evaluation of our shape recognition methods. We aim at automatic classification of petroglyph shapes into predefined classes (see Figure 2.15). Typologies for petroglyphs are strongly related to the site. Hence, each site or local group of sites has its own typology. For our ground truth, we used two typologies. One typology was developed by a small group of 3D-Pitoti project members (amongst them the author). The second typology is based on the typology used in the most recent large-scale site publication about the Campanine site in Valcamonica by Umberto Sansoni [SG09]. The Campanine site contains - in contrast to the rest of Valcamonica - many medieval figures. Hence, it was necessary to extend the typology in order to be appropriate for all rock art sites in our dataset. This extension was done by Craig Alexander. We use the annotated shapes in three chapters. In Chapter 7 we use a small set of 100 Pitoti to evaluate a shape descriptor based on the skeletal graph. In Chapter 8 we refine the skeletonisation on a set of more than 1000 Pitoti. In Chapter 9 we evaluate our combination of shape descriptors on a set containing more than 1000 Pitoti. These 1000 Pitoti are classified into the two typologies we developed for the project. In the following paragraph we describe the properties of the full dataset used for shape classification.

The dataset consists of 1388 distinct petroglyphs (see Figure 2.15 for a part of the material). For the experimental investigation of the classification capabilities in

2. MATERIAL



Figure 2.15: Petroglyph shapes from our database. The underlying tracings are ©CCSP and Alberto Marretta, used with permission.

Subsection 9.2.1 we need to form classes that serve as ground truth for the evaluation of our classifiers. We aim at using the whole dataset in our experimental evaluation. Hence, the straightforward approach is to use each endpoint of a typology-tree branch as a class. However, to provide a minimum amount of training data for a class, we only use branch endpoints where at least 10 examples exist in our dataset. The two typologies provide different views on the dataset. Thus, they have different branch endpoints with different numbers of examples. Hence, the total numbers of petroglyphs investigated for the two typologies are slightly different. See Tables 2.2 and 2.1 for the classes used.

2.6.3 Shape Annotation Tool

We developed an interactive tool that allows the precise annotation of digitised classic petroglyph tracings. The tool supports the user in the manual classification of figures and provides all necessary annotation fields specified by the two typologies introduced in Subsection 2.6.2. The figures are annotated according to both typologies in parallel. The exact shape of a figure and the textual annotation data is stored in a database on a central server to facilitate further processing. The annotation tool is realised as a web application that is accessible via the internet with any modern web browser. Thus it can be used from any computer, independent of the operating system.

2.6.3.1 Main Features

The tool makes use of several features. We developed some of the features customarily and adopted others from open source projects to meet our needs.

CanvasZoom: For good user experience with acceptable performance when navigating within the high resolution tracings, we use a tile-zooming algorithm with different zoom levels, comparable to web-maps like Google Maps.

Different selection tools: To meet the different requirements for selecting the various types of traced figures, we implemented a polygon selection tool, a rectangle selection tool and a magic wand selection tool.

Special filiform selector: For selecting filiforms (scratched line-like figures, see the arrow and the bow in Figure 2.14(d) for an example) we provide an alternative selection mode that uses a different underlying colour model (LAB instead of RGB).

Classification: Drop-down menus allow the classification of a selected figure into two different typologies. If the membership in a particular class is not certain, a confidence value between zero and one can be added to the classification.

Multi-user functionality: To improve the level of confidence in the annotation of figures, multiple users can individually annotate each figure. An administrative user can see and compare all annotations.

Statistical reports: To allow a progress overview various statistical reports can be generated from the database.

2. MATERIAL

Table 2.1: Classes based on branch endpoints of the modified Sansoni typology.

ID	Type	N
	Prehistoric / Protohistoric - Natural world - Anthropomorph: armed	
1	- Armed figure with shield and spear (no helmet)	61
2	- Armed figure with shield and sword (no helmet)	105
3	- Armed figure with sword (no shield or helmet)	10
4	- Armed figure with shield only	10
5	- Armed horseman	39
	Prehistoric / Protohistoric - Natural world - Anthropomorph: unarmed	
6	- Unarmed horseman	9
7	- Oranti praying figure	77
8	- Other anthropomorph not elsewhere classified	21
	Prehistoric / Protohistoric - Natural world - Animal	
9	- Bird	57
10	- Deer	60
11	- Dog	60
12	- Horse	56
13	- Other animal natural	77
	Prehistoric / Protohistoric - Material culture - Structure	
14	- Hut	143
	Prehistoric / Protohistoric - Material culture - Other material culture	
15	- Palette	35
	Prehistoric / Protohistoric - Material culture - Weapon (not wielded)	
16	- Axe	43
	Prehistoric / Protohistoric - Symbolism - Writing	
17	- North etruscan inscription	11
	Prehistoric / Protohistoric - Symbolism - Topographic form	
18	- Rectilinear filled form with line parallel to one long side	15
19	- Rectilinear filled form without line parallel to one long side	14
	Prehistoric / Protohistoric - Symbolism - Specific symbol	
20	- Rosa camuna	10
21	Prehistoric / Protohistoric - Other unclassified	9
	Medieval and post-medieval - Material culture - Other material culture	
22	- Key	39
	Medieval and post-medieval - Symbolism - Cross	
23	- Cross potent - cross with blocks at the terminals	30
24	- Greek cross equal width/height	19
25	- Latin cross taller than wide	36
	Sansoni not securely datable - Geometrical form - Planar symbol	
26	- Star	20
Σ		1066

Table 2.2: Classes based on branch endpoints of the 3D-Pitoti project typology.

ID	Type	N
	Inscription	
1	- Local variation of the Etruscan alphabet	11
	Abstract	
2	- Camunian rose	10
3	- Map	56
4	- Star	20
5	- Circle	14
	Object	
6	- Cross	106
7	- Key	39
8	- Palette	35
9	- Shoe/footprint	98
10	- House	143
	- Weapon	
11	* Axe	43
	Animal	
12	- Dog	60
13	- Deer	60
14	- Horse	55
15	- Bird	57
16	- Other	85
17	Anthropomorph	335
	<i>Please refer to Figure 9.6 for the morphological subdivisions of the Antropomorph class.</i>	
Σ		1227

2.6.3.2 Handling and Annotation Workflow

After logging in, the user first has to select a site and then a tracing of a rock panel within which he wants to classify figures. Subsequently, the user is presented with an overview of the whole tracing (Fig. 2.16). Already annotated figures on a tracing are indicated with a grey bounding box. With a click on that bounding box, the exact shape of that figure and its classification data are loaded from the database (Fig. 2.17). The annotation process of a new figure is divided into several steps.

Zoom: The user can use the mouse wheel or the zoom-controls (buttons/slider) in the user interface to zoom to maximum zoom level. Only at the maximum zoom level a detailed shape selection can be made.

Selection of a Figure: To navigate within the image, the user can use the hand-tool to click and drag the tracing. To select the shape of a figure, the user has a choice amongst three tools. With the rectangle selection tool, the user first defines a rectangular area of the tracing and subsequently indicates a colour by clicking on a tracing part within the previously selected rectangular area. All pixels within the rectangle which have the same colour as the clicked tracing part are selected and highlighted in yellow. The polygon selection tool has the same functionality except that the area selected in the first step

2. MATERIAL

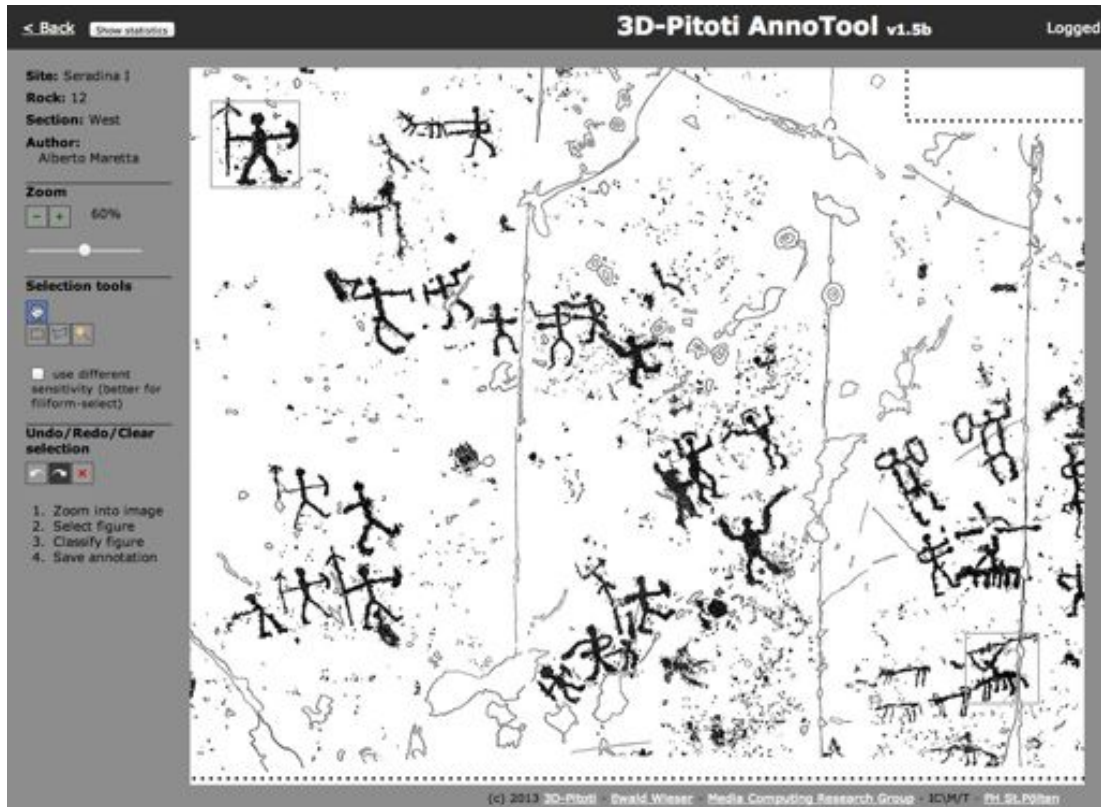


Figure 2.16: Overview of the ground truth annotation web tool interface.



Figure 2.17: View of a figure in its rock panel tracing.

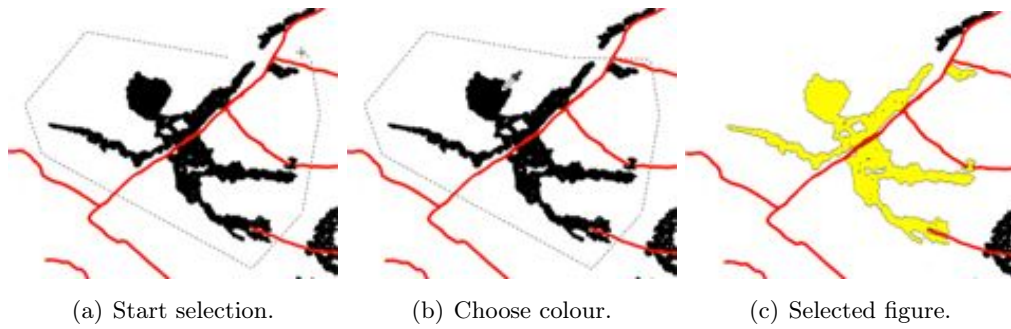


Figure 2.18: Use of the polygon selection tool.

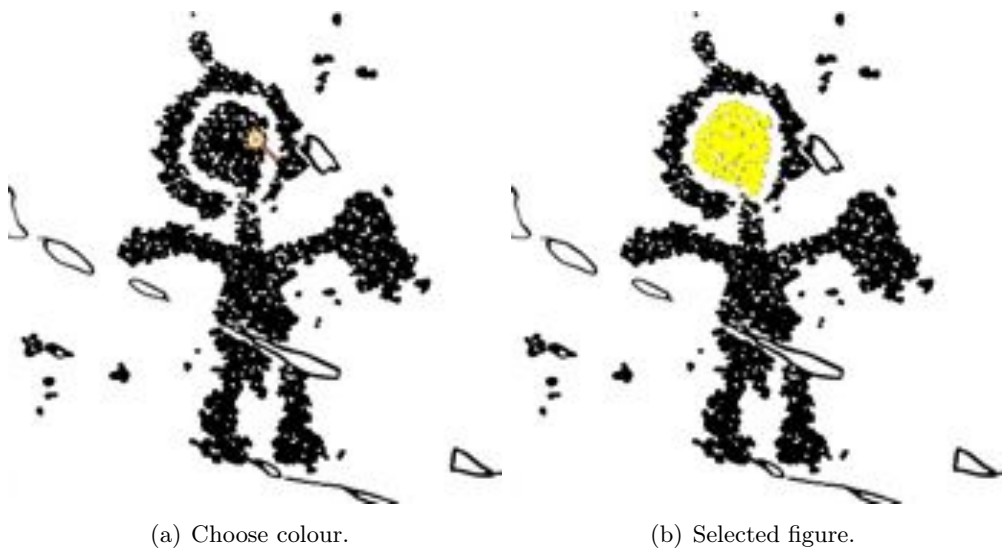


Figure 2.19: Use of the magic wand selection tool.

2. MATERIAL

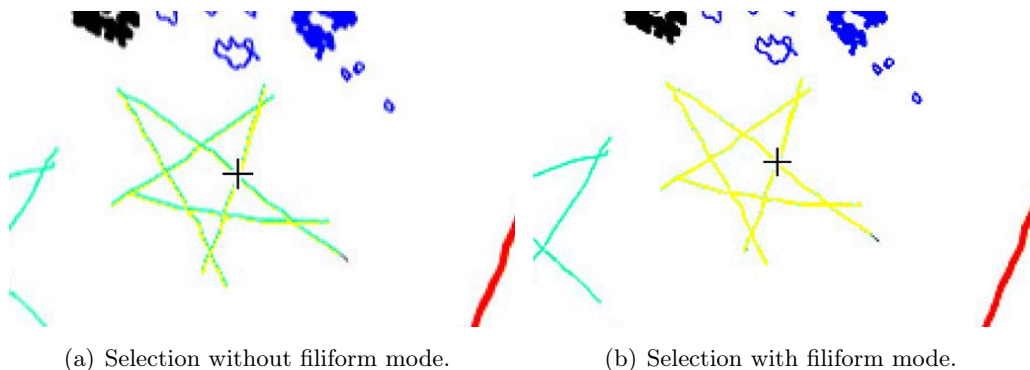


Figure 2.20: Comparison of selection without and with filiform selection mode.

can be a polygon (Fig. 2.18). The magic wand selection tool allows the user to select connected pixels with the chosen colour. Thus, large figures with low fragmentation can be selected easily (Fig. 2.19). A special filiform selection mode enables the user to select line-like figures. This mode utilizes another colour model for all selection tools (LAB instead of RGB), which facilitates the correct selection of all pixels of the line-like figure (Fig. 2.20).

Existing selections can be extended or reduced through addition or subtraction of selections. If the user makes a mistake, the tool allows up to three steps of undo and redo.

Classification of a Figure: As soon as the user has finished the selection of the shape of a figure he can start the classification process. At the beginning he is presented with one drop-down menu for the main classes of each of the two typologies. After selecting a figure class in the drop-down menu, another drop-down menu with the classes from the next level of the typology is revealed (Fig. 2.21). An item from each level of drop-down boxes must be chosen in order to complete the classification process. If the user is unsure about the classification of a figure, he can reduce the confidence value with a slider and add a second classification set for the figure within the same typology. The sum of the individual confidence values has to be 1 in order to complete the classification process.

Special cases (superimposition, etc.) that might occur when selecting a figure (Fig. 2.22), are marked with flags which can be set in the classification process (Fig. 2.21).

2.6.3.3 Statistical Reporting

The user can access a list of all figure classes that contains the number of figures of each class that he has already annotated (Fig. 2.23). Besides that, he can see how many figures of each class are annotated in total, including the figures annotated by all other users. The user can then concentrate on the annotation of certain figures, to populate the ground truth dataset according to pre-established targets.

With a click on the number of figures in each class, the user gets an overview of thumbnails of all images in that class (Fig. 2.24). With a further click on one of the

Figure 2.21: Classification process.

thumbnails, the user is presented with a detail view of that figure including its location (site, rock and section), the full image and the classification data (Fig. 2.25). He can edit the classification data or delete the figure from the database. A click on the image or on the “*Show figure in tracing*” link opens the view of the figure in the context of the rock panel tracing to which it belongs (Fig. 2.17). If necessary, the selection of the shape of a figure can be modified using the selection tools.

2.7 Conclusion

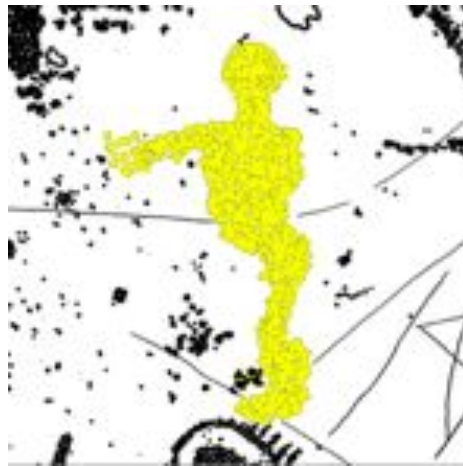
In this chapter, we have introduced our material and its origin. We described rock art in general, Valcamonica (the site where our material stems from) and several archaeological aspects in Sections 2.1 to 2.3. We observe, that the methods we will develop in this thesis may support archaeology in two aspects, namely *documentation* and *analysis*. The automated segmentation of rock surfaces in pecked and natural regions as well as the subsequent classification of figures have the potential to substitute classic documentation methods. The high-resolution 3D scans allow analyses of pecking style that are not possible with the naked eye.

To evaluate our methods, we need data for each task and respective ground truth sets that are annotated by experts. In Sections 2.4 and 2.5 we described the 2D images and the 3D scans we will use for segmentation and pecking style investigation. Section

2. MATERIAL



(a) Figures with superimposition.



(b) Figure not completed by artist.

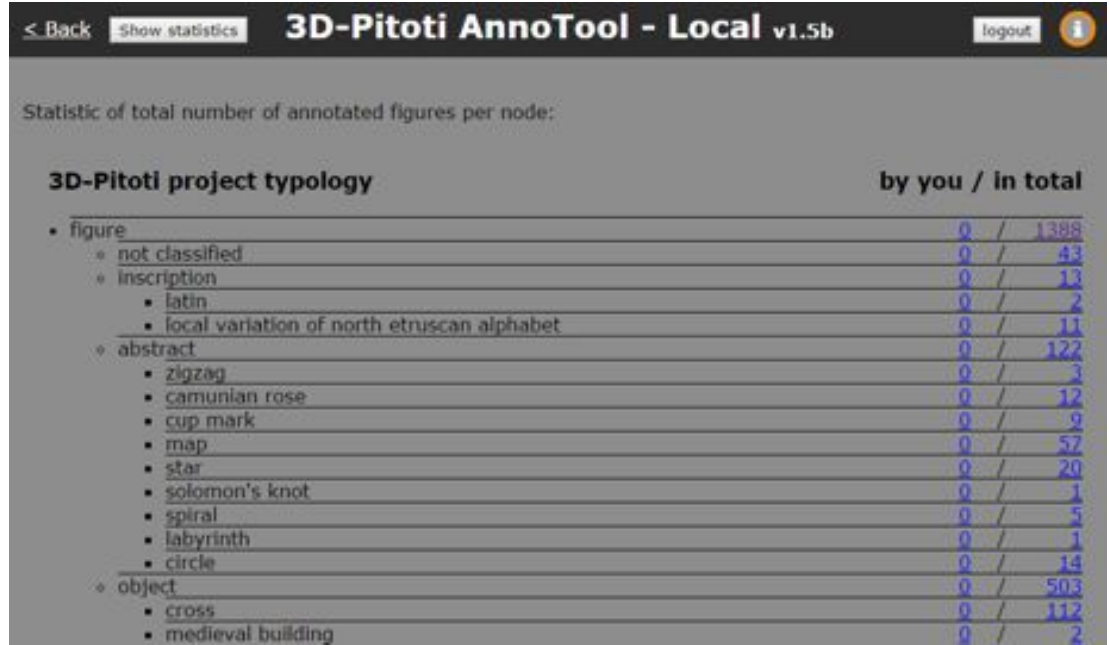


(c) Figure damaged.



(d) Tracing incomplete.

Figure 2.22: Special cases when selecting a figure.



Statistic of total number of annotated figures per node:

3D-Pitoti project typology	by you	in total
• figure	0	1388
◦ not classified	0	43
◦ inscription	0	13
▪ latin	0	2
▪ local variation of north etruscan alphabet	0	11
◦ abstract	0	122
▪ zigzag	0	3
▪ camunian rose	0	12
▪ cup mark	0	2
▪ map	0	57
▪ star	0	20
▪ solomon's knot	0	1
▪ spiral	0	5
▪ labyrinth	0	1
▪ circle	0	14
◦ object	0	503
▪ cross	0	112
▪ medieval building	0	2

Figure 2.23: Statistics of numbers of figures in each class. Please note, that the screenshot was taken at a later date than our experiments. As annotation is an ongoing process, the numbers are higher.

2.6 deals with our ground truth data. In Subsection 2.6.1 we described the ground truth annotation for 2D and 3D segmentation that resulted in binary masks encoding pecked rock surface and natural rock surface. In Subsection 2.6.2 we presented the ground truth annotation for the shape of petroglyph figures. Due to the limited amount of 3D scans and consequently limited amount of petroglyph shapes we need a proxy dataset. We use 2D images of classical petroglyph tracings. In Subsection 2.6.3 we presented a web tool that allows the large-scale annotation of petroglyph shapes in classic tracings. More than thousand petroglyph shapes have been annotated by experts. We observe, that significant efforts were made to establish the datasets required for this thesis. A large share has been achieved within the 3D-Pitoti project (see Chapter 1).

2. MATERIAL

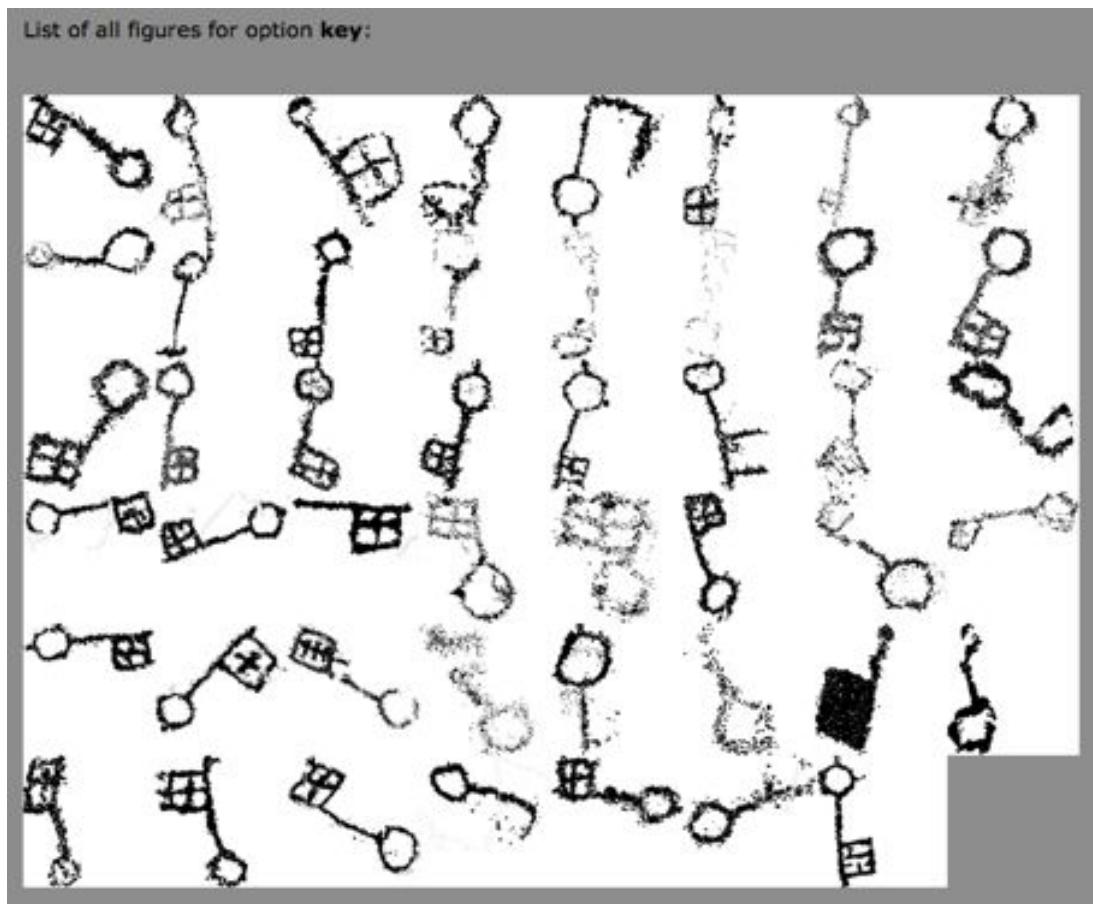



Figure 2.24: All figures for class *key*.

Details for figure 331:

Site: Campanine
Rock: 7
Section: A



classified by: Giovanna
on: 2014-01-15 14:51:40

[Show figure in tracing](#)

3D-Pitoti project typology

Classificationset:
Select figure:
object
Select object:
key
Set confidence for above classification:

Total confidence: 1

Sansoni typology - Alexander modification

Classificationset:
Select image period:
medieval and post-medieval
Select medieval and post-medieval:
material culture
Select material culture:
other material culture
Select other material culture:
key
Set confidence for above classification:

Total confidence: 1

Figure contains superimposition:
Figure is incomplete due to author:
Figure is damaged:
Figure is incomplete due to missing parts of tracing:

Figure 2.25: Details for one *key* figure.

2. MATERIAL

Petroglyph-related Work



Figure 3.1: A petroglyph depicting a deer in Valcamonica, Italy.

THIS chapter contains work related directly to rock art and petroglyphs which is relevant for our tasks. Please note that for better readability, the technical work related to the different analysis steps is located in the respective chapters of this thesis. First, we look into 3D scanning of rock art as an important part of our material stems from 3D scans. Second, we present related work on the segmentation of Petroglyphs. Third, we present approaches that investigate pecking styles. Finally, we present work related to shape classification of petroglyphs.

3. PETROGLYPH-RELATED WORK

3.1 3D Scanning of Petroglyphs

3D scanning has been used in archaeological documentation since more than a decade now. A well-known early example is Stanford’s approach to scan several Michelangelo statues in Florence, Italy by Levoy et al. [LPC⁺00]. In the “Digital Michelangelo” project, they built a system containing hardware (laser range finders, digital cameras) and software to acquire high precision 3D outside the laboratory. The largest model they acquired is Michelangelo’s David with (at the time of the scanning expected) 2 billion polygons and 7000 color images. It took years to reconstruct 3D models with high accuracy. The project’s website states in news from 2009, which is 9 years (!) after the acquisition of the scans: “We now have a full-resolution (1/4mm) 3D model of Michelangelo’s 5-meter statue of David. The model contains about 1 billion polygons.”¹ Scopigno et al. summarize the possible uses of 3D scans of cultural heritage artifacts “beyond plain visualisation” [SCC⁺11]. They state, that 3D is used in cultural heritage for two major tasks, a) studying artwork and b) serving as a support medium for indexed archival knowledge. The examples include the detailed study of artworks, documentation for on-site archaeology, restoration of fragmented artwork and virtual restoration of painted decorations.

An early example in the rock-art domain is presented by Landon and Seales who propose a system for 3D acquisition and presentation of Puerto Rican petroglyphs [LS06]. They consider a full pipeline from acquisition over scholarly use of the scans to presentation for the general public. They use a laser-based structured light system (SLS) to acquire the rock surfaces. They describe a software which allows interactive surface manipulation (e.g. false lighting), surface mark-up and large-scale display. Díaz-Andreu et al. use laser scanning and photogrammetry to scan a rock surface in Cumbria on which two decades earlier a spiral had been documented [DABRR06]. Although the spiral could not be found again, the authors state that the quality of the 3D documentation is far beyond traditional documentation regarding precision and accuracy. Lerma et al. describe the use of terrestrial laser scanning (TLS) for the documentation of the complete cave of Parpalló (Spain) with peckings dating back to the Upper Paleolithic era [LNCV10]. Additionally to the TLS they use close range photogrammetry for the pecked parts of the surface. Gonzalez-Aguilera follow a comparable approach in Pindal cave in Spain [GAMNRGM11]. Sanz et al. use a standard digital camera and photogrammetry to record petroglyphs at several sites in Galicia (Spain) [SDR⁺10].

A recent approach that considers multiple scales to acquire 3D data of petroglyphs together with the surrounding rock panel has been proposed by Alexander et al. [APR15] in the 3D-Pitoti project. The authors describe the necessities for different scales of acquisition for a holistic approach to acquire 3D data of archaeological sites. They capture the sites together with the surrounding landscape in three scales - macro (meters), mid-range (centimeters) and micro (sub-millimeter, see Figure 3.4) and describe methods for the mid-range and the micro scale. They use un-manned aerial vehicles (UAVs) for the mid-range and a specially built “rock art scanner” based on

¹<http://graphics.stanford.edu/projects/mich/>



Figure 3.2: The 3D-Pitoti project scanner, ©Axel Pinz, used with permission.

3. PETROGLYPH-RELATED WORK

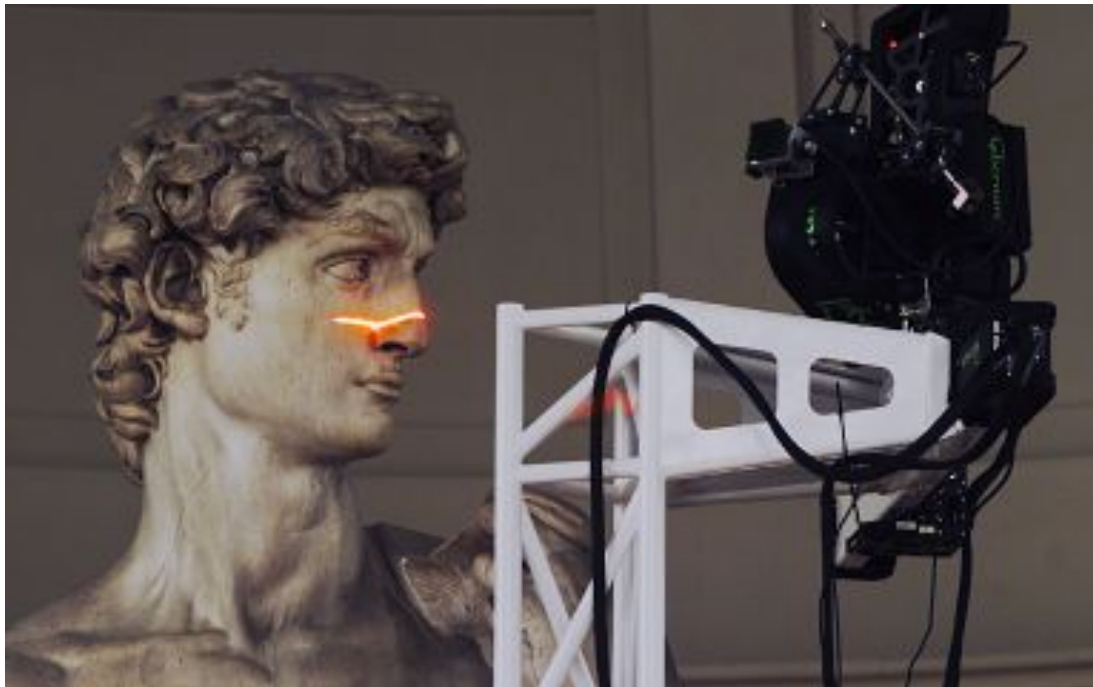


Figure 3.3: An early example of high-resolution 3D acquisition of a rock surface. Laser scanning of Michelangelo’s David.¹

photogrammetry for the micro-scale (see Figure 3.2). In this thesis we use the ground truth data of the micro-scale method (see Appendix A).

3.2 Segmentation of Petroglyph Images

The next step after acquisition is segmentation of the rock surface in pecked and natural surface areas using digital images or 3D scans of the surface. The automatic segmentation of rock surfaces containing petroglyphs has rarely been addressed so far. Zhu et al. consider the automated segmentation using digital images infeasible, hence they propose a collaborative manual segmentation approach that utilizes CAPTCHAs for rock art image segmentation [ZWK10]. Deufemia and Paolino apply segmentation to digitized manual tracings [DP14]. This is a rather straight-forward task since manual tracings are binary images that already represent a manual pre-segmentation of the figures.

¹©Digital Michelangelo Project, Stanford University, used with permission.

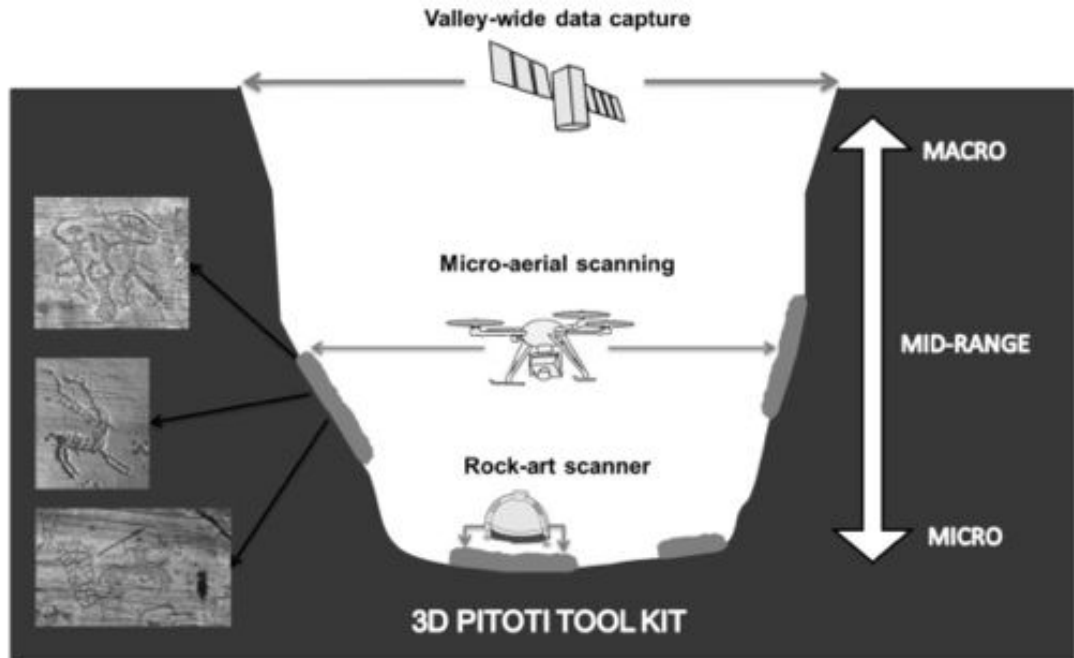


Figure 3.4: Scales of acquisition for the holistic approach of capturing a complete valley in the 3D-Pitoti project.

3.3 Pecking Style Similarity

Zotkina et al. conduct an experimental study of percussion and tool types in the Minusinsk basin in Siberia [ZTKP14]. They replicate different peck marks with a diverse set of tools. Moitinho et al. approach a numeric analysis of pecking styles using terminology and software from geomorphology [dARBP13]. They aim to estimate the tools and gestures used for creating rock art. The approach includes experimental archaeology to replicate ancient peck marks. They scan the surfaces with 0.28mm resolution. From the scans, they use patches with $50 \times 50 \text{mm}^2$ containing homogeneous peck marks to extract surface metrology parameters with proprietary software. They cluster the numeric descriptions of the peck marks and observe three clusters which contain surfaces with the following properties: 1) fine, 2) coarse and regular, 3) coarse and irregular. They do not observe clusters linked to the manufacturing process of the samples.

3.4 Shape Classification of Petroglyphs

Only a small number of papers has been published in the field of automated petroglyph classification. Other areas of archaeology in which automatic image classification has been investigated range from finding archaeological sites in aerial photographs [TLS09, Kva13] to identify individual characters in monumental Maya inscriptions [RRPOGP11].

3. PETROGLYPH-RELATED WORK

The latter application is perhaps closest to what our approach and the related work discussed below aim to achieve: Maya glyphs exhibit considerable variation in the way a given character is carved, much like the variation one finds in, say, the petroglyph depiction of a deer in Valcamonica.

Numerous surveys about shape analysis have been published, comprehensive and important in this field are the surveys by Pavlidis [Pav78], Loncaric [Lon98], Zhang and Lu [ZL04] and Yang et al. [YKR08]. Classifications and taxonomies of shape descriptors have been proposed in the mentioned surveys in different variants. A widely used common denominator is the distinction between contour-based and region-based descriptors. We observe that the petroglyph shapes in our dataset are perceptually close to skeletons. Hence, we add a third category of skeleton-based approaches and summarize region-based, contour-based and skeleton-based petroglyph-related approaches.

Region-based descriptors: Zhu et al. propose the use of a slight modification of the Generalised Hough Transform (GHT) for the mining of large petroglyph datasets [ZWKL11]. The main arguments for GHT and against other shape similarity measures are the existence of petroglyph images where a single petroglyph consists of several parts and the possibility of merged parts of petroglyphs that drastically change the topology of the petroglyphs. They extensively evaluate their approach and achieve good results. However, they mostly evaluate their approach on proxy datasets (Farsi digits, sketches) and synthetic petroglyph shapes¹. Deufemia et al. use the radon transform of petroglyph images as the shape descriptor for unsupervised recognition via self-organizing maps (SOM) [DPdL12]. In a second step, they use a fuzzy visual language parser to solve ambiguous interpretations by incorporating archaeological knowledge. They evaluate the approach on a large dataset and achieve good results. Deufemia et al. use an Enhanced version of Generic Fourier Descriptors (EGFD) to detect and classify petroglyphs from full scenes [DP14]. In their experiments on 1215 petroglyphs in 53 scenes EGFD outperforms the descriptors GHT and Shape Context. They do not combine the descriptors.

Contour-based descriptors: Deufemia et al. propose a two stage classification of petroglyphs [DP13]. They use shape context descriptors [BMP02] to provide an initial raw clustering with self-organizing maps. In the second step, they use an image deformation model to classify the petroglyph shapes. They evaluate their approach on a relatively small dataset (17 classes with 3 examples each) that they enlarge by using 30 affine transformations of each image.

Skeletal descriptors: To our knowledge, we are the first to utilize skeletal descriptors for petroglyph classification.

3.5 Discussion

We discuss related work for segmentation, pecking style similarity and shape classification as we follow these topics throughout the thesis.

¹See <http://alumni.cs.ucr.edu/~qzhu/petro.html> for the evaluation data.

Segmentation: Prior to our work, automatic segmentation has been considered infeasible. The existing approaches are either manual [ZWKL10] or work on material that is already a binary image [DP14]. To our knowledge, we are the first to approach automatic segmentation based on digital photographs as well as on 3D point clouds.

Pecking Style Similarity: In contrast to Moitinho et al. we approach the peck mark similarity estimation with 3D surface descriptors which have been developed for registration and object detection. To our knowledge, there is no work on automated peck mark similarity estimation comparable to ours.

Shape Classification: All existing approaches dealing with petroglyphs have drawbacks. Zhu et al. evaluate at large scale on publicly available data, but the data are different from the material we are using [ZWKL11]. They use *inter alia* sketch data and synthetic petroglyph shapes. In contrast to that, our data is the product of the standard archaeological documentation process which is manual tracing of individual peck-marks on transparent sheets (see Subsection 2.6.1). Deufemia et al. evaluate on small datasets [DPdL12][DP13]. In an other approach they have a large dataset and compare three different descriptors in their evaluation, but they don't evaluate the combination of descriptors [DP14]. Moreover, they only compare region-based descriptors. In our approach, we aim to overcome these shortcomings.

3. PETROGLYPH-RELATED WORK

Segmentation of Petroglyph Images

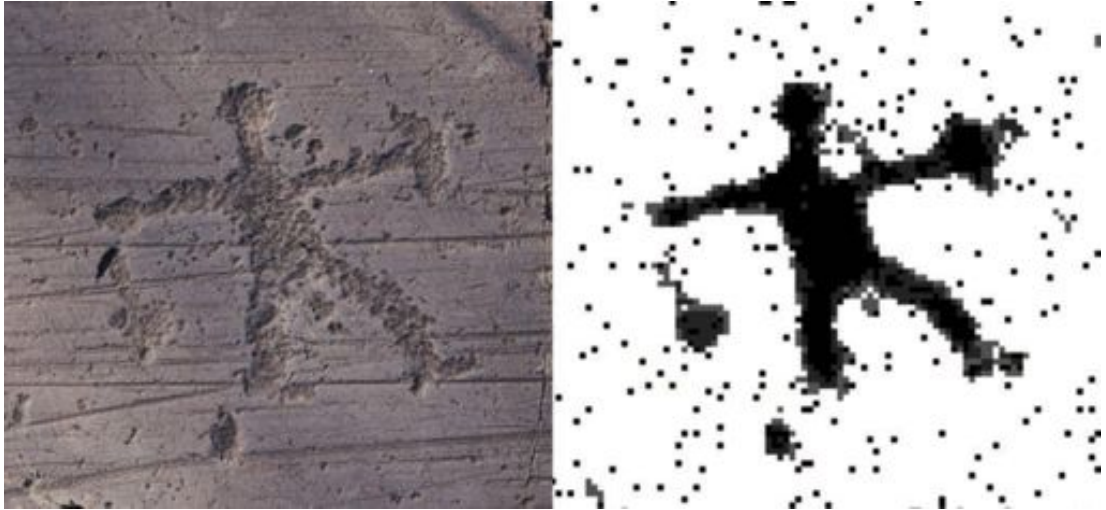


Figure 4.1: A human figure and its automatic segmentation result.

DIFFERENT time-consuming manual methods to determine and document the exact shapes and spatial locations of petroglyphs on one rock panel have been carried out over the past decades [AC94]. In this chapter, we address our goal to support the documentation work with (semi-)automatic tools by an approach that uses high-resolution photographs of rock surfaces. In Chapter 5 we will investigate segmentation based on 3D point clouds. In both chapters, we approach segmentation of petroglyphs as *pixel- or point-wise classification*.

Segmentation of petroglyphs based on photographs of the rock surface (see an example in Figure 4.1) is a challenging task. Figure 4.2 shows a detail of our rock art material (for more details on our material please refer to Section 2.4). The depiction is highly dependent on the direction of the light. Glacial polish causes noise in the non-pecked regions. The exact boundaries of the pecked regions are hard to determine, as the pecks are of varying depth and quality. Abrasion and alteration change the appearance of the rock and the artwork.

We approach segmentation as pixel classification. We propose novel feature variants

4. SEGMENTATION OF PETROGLYPH IMAGES



Figure 4.2: Detail of rock art material.

and evaluate visual features, feature variants and feature combinations. We show that with the appropriate combination of different visual features, good segmentation results are possible. Furthermore, we demonstrate that varying a simple threshold for classifier fusion, the user can interactively change the position of the result on the recall-precision curve.

In Section 4.1 we present related work. Section 4.2 describes our approach. In Section 4.3 we describe the evaluation of our approach and in Section 4.4 we present the results.

4.1 Related Work

We summarize related work in fields with pixel-wise classification approaches comparable to our task and in texture classification. Yin et al. [YLH⁺09] use color and edge features in a k-NN classifier for rock structure classification in Formation MicroImager (FMI) images¹. Partio et al. [PCGV02] use gray level co-occurrence matrices (GLCM, see [HSD73]) and Gabor filters to model textures of rock images. They perform classification with a k-NN classifier. The results of the evaluation on a test database containing 168 rock images in 7 classes are good for rock images that have almost homogeneous textures. GLCM performs better than Gabor filters. The authors suggest that prior texture segmentation should be performed for rock images with multiple textures.

Khoo et al. [KOW08] model textures as GLCM and use a support vector machine (SVM) to classify textures. They evaluate their segmentation approach on few synthetic texture mosaics and two satellite images with good results. Kim et al. [KJPK02] use a support vector machine (SVM) for texture classification. They use the pixel intensities as input for the SVM, i.e. no prior feature extraction is performed. The evaluation of their approach against synthetic texture mosaics delivers good results. Varma and Zisserman [VZ05][VZ03] use textons (see [LM99]) as texture models. They evaluate their approach on the Columbia-Utrecht reflectance and texture database (CURet [DvGNK99]) and achieve very good classification results with and without the usage of filter banks.

4.2 Approach

4.2.1 Pixel Classification

We define any pixel that is inside a petroglyph as foreground pixel and subsequently any other pixel as background pixel. First, for each pixel to classify we obtain a local window of the input image with this pixel in the center. Second, we extract visual features of each of these local windows. Third, we train an SVM. Fourth, we optionally

¹The Formation MicroImager is a proprietary micro-resistivity imaging device used for borehole imaging, see http://www.slb.com/~media/Files/evaluation/brochures/wireline_open_hole/geology/fmi-hd_br.pdf.

4. SEGMENTATION OF PETROGLYPH IMAGES

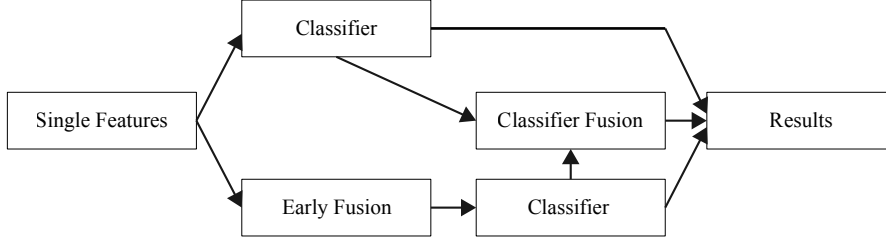


Figure 4.3: Schema of different fusion strategies applied.

fuse several classifiers (see Section 4.2.2). Fifth, we classify new data with the models obtained in step three or step four.

Features for object classification include color, edge and texture based features. Datta et al. [DJLW08] state, that the major types of features in image retrieval are color, texture, shape and salient points. Shape features are not suitable for our task, as shape is an attribute of an image segment, i.e. shape features are extracted post segmentation. Since our material contains too many salient points (i.e. interest points or corner points) due to the structured surface (see Subsection 4.3.1) we rule out salient points as well. The three feature categories we consider for our task are therefore color, edge and texture.

4.2.2 Fusion Strategies

The different features incorporated in our pixel classification approach may contain complementary information. Therefore, we apply different combinations of these features with different fusion strategies. The strategies we employ are *Early Fusion* and *Late Fusion* or *Classifier Fusion* (see Figure 4.3).

Early Fusion: We combine the feature vectors of different features by concatenation. In the next step, we use the combined feature vectors to train the classifier.

Classifier Fusion: We utilize the output labels from single feature classifiers and early fusion classifiers. We fuse the output labels. Numerous methods to fuse classifier output labels have been proposed and evaluated. A frequently used strategy is majority voting. Consensus can be achieved by unanimity (all agree), simple majority (50%+1 agree) or by plurality (most votes) [Kun04]. In our two-class case, majority is equivalent to plurality. To fuse our classifier outputs, we propose an extended version of majority voting. We utilize the number of votes for a foreground pixel in two ways. First, we assume that the number of votes is equivalent to the confidence in the fusion result for this pixel. In the output image of the classification results, we color foreground pixels with high confidence, medium confidence and low confidence differently. Second, we assume that the best segmentation output of our classifier in terms of used measures is not necessarily the output that suits our needs best.

The sizes of our two classes (foreground and background pixel) differ by one order of magnitude. Hence, we measure the quality of our results with recall, precision and

$F1$, where tp is the number of true positives, fp is the number of false positives and fn is the number of false negatives.

$$precision = \frac{tp}{tp + fp} \quad (4.1)$$

$$recall = \frac{tp}{tp + fn} \quad (4.2)$$

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (4.3)$$

A use case for segmentation results is the highlighting of pecked regions on an image of a rock panel. An optimization for precision could lead to better use of results.

Thus, we use a modified majority voting. The user changes a threshold between a value of 1 and the number of classifiers to fuse. The value of this threshold corresponds with the number of votes needed to classify a pixel as foreground pixel. If the threshold is 1, a pixel is classified as foreground pixel if any of the fused classifiers classified this pixel as foreground pixel. Subsequently, the number of votes needed for the majority increases with the threshold. If the threshold is at its maximum, a pixel is classified as foreground pixel, if all fused classifiers labelled this pixel as foreground pixel.

4.3 Experiments

We extract the visual features out of local windows with a size of 128x128px. We chose the window size based on preliminary experiments with window side lengths of 32, 64, 128 and 256px. 128px was the best trade-off between distinctive performance, computational demand and blurring. We extract color, edge and texture features (see Subsection 4.3.2) out of the windows with a horizontal and vertical stepsize of 16px. We use a support vector machine (SVM) for classification. We train the classifier with single features and employ several fusion strategies (see Subsection 4.2.2). For validation, we utilize repeated random sub-sampling with 10 sub-samples [Rao05].

4.3.1 Material

We use two carefully selected 21 MPixel images of rock surfaces containing petroglyphs. They are differently lit and the typical rock structure appears visually different in the two images. Furthermore, they contain petroglyphs with different pecking styles. The test image composed of these two images has a size of more than 40 million pixels. We use the images in full resolution. The material poses several challenges. Grass is growing in cracks in the rocks. The surfaces of the rocks have scratches and other traces of abrasion and alteration (see Figure 4.4). The petroglyphs are pecked with different styles, i.e. with different tools by different authors. For details about the employed material please refer to Section 2.4.

4. SEGMENTATION OF PETROGLYPH IMAGES



Figure 4.4: Small details of our test images (see Figures 2.9(a) and 2.10). Problems of the material include grass (tl), shadows (tr), scratches (bl) and cracks (br).

Table 4.1: Features extracted from each local window.

Feature Abbreviation	Feature Description
RGBHist M	RGB histogram with M bins
LumHist N	Luminance histogram with N bins
DSIFT_gauss	Dense SIFT with prior gaussian blurring
DSIFT_nogauss	Dense SIFT without prior gaussian blurring
GLCM O	Gray level co-occurrence matrix of order 2 calculated from O gray levels
GLCMStat O	21 statistical measures calculated from GLCM O
GLCMPOrder Q	Gray level co-occurrence matrix of order Q calculated from P gray levels
GLCMStat P Order Q	21 statistical measures calculated from GLCMPOrder Q
LBP_Hist_u2_n U _r V	Local binary pattern histogram with U sampling points on a radius of V

4.3.2 Visual Features

Table 4.1 shows an overview of the extracted features. We describe some of the extracted features in more detail.

Dense SIFT: Lowe [Low04] proposed the widely used SIFT features. We do not need the prior salient detector, and we do not need rotation or scale invariance, thus we extract dense SIFT features with and without prior Gaussian blurring [VF08].

Gray Level Co-Occurrence Matrix: Haralick et al. proposed the classic feature to describe texture in 1973 and evaluated it for crop classification [HSD73]. The image is scaled to N gray-levels. The gray levels of neighboring pixel-pairs are compared and each occurrence of a distinct combination is counted. This results in a matrix of size $N \times N$ containing these counts. Haralick proposed several statistic measures for this matrix that are used as feature vector. GLCM and GLCMStat for order 2, i.e. the comparison of two pixels, are widely used visual standard features. But there are hardly any evaluations of higher order GLCMs (e.g. by Anys and He 1995 [AH95]). The memory requirements of higher order GLCMs grows exponentially. A GLCM of order 3 (comparison of three pixels) results in a matrix of size $N \times N \times N$. We had a look at GLCMs of order 3 in our evaluation. $N=8$ results in 512 feature dimensions, $N=16$ results in 4096 feature dimensions. The Matrices are sparse, to reduce dimensionality we discarded the sparsest feature dimensions in the training data set and subsequently the same dimensions in the complete data set. Inspired by the measures Haralick proposed, we calculate 21 statistical measures from each feature vector.

Local Binary Patterns: Ojala et al. proposed local binary patterns (LBP) as multiresolution gray-scale and rotation invariant texture features in 2002 [OPM02]. LBPs are widely used in several application fields, e.g. crop classification and face recognition.

4. SEGMENTATION OF PETROGLYPH IMAGES

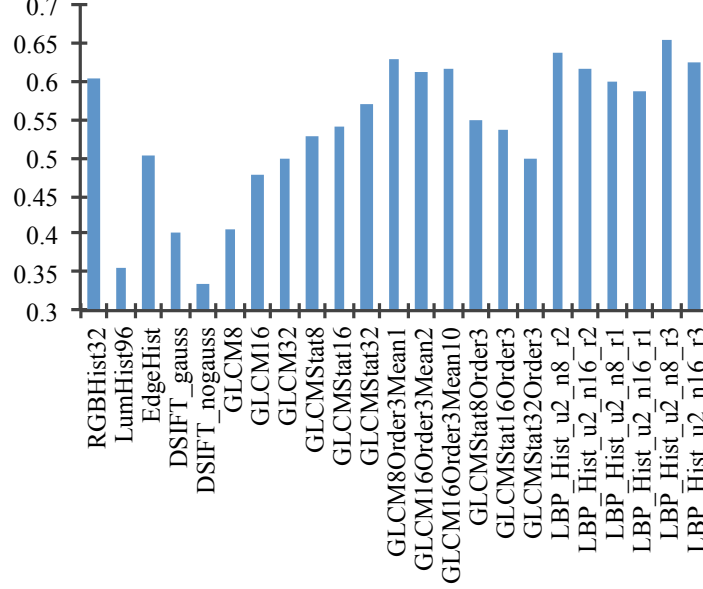


Figure 4.5: F1- results of single feature classifiers.

4.3.3 Training and Classification

Single Features: We train a classifier with each single feature. We evaluate the performance of the models to learn about the isolated distinctiveness of each feature.

Early Fusion: Based on the results of single feature classifiers, we combine different features by concatenation. We assume that different single feature categories contain complementary information. We use the concatenated feature vectors to train the classifier. We evaluate the performance of the early fusion models.

Late Fusion: We combine the output labels of different classifiers with the strategies explained in Subsection 4.2.2. We fuse the classifiers with the best results employing majority voting. We utilize single feature classifiers as well as early fusion classifiers. We assume that even early fusion classifiers with nearly similar features and results contain complementary information. We want to clarify, whether late fusion improves the results compared to early fusion.

Late Fusion with Thresholded Majority Voting: We use the best result from the previous paragraph and vary the voting threshold. We visualize the results to answer the following question: Does a combination of the different late fusions improve the segmentation?

Interactive Late Fusion: The user can change the threshold for the majority voting interactively (see Subsection 4.2.2). We use the same classifier combinations as in the previous paragraph.

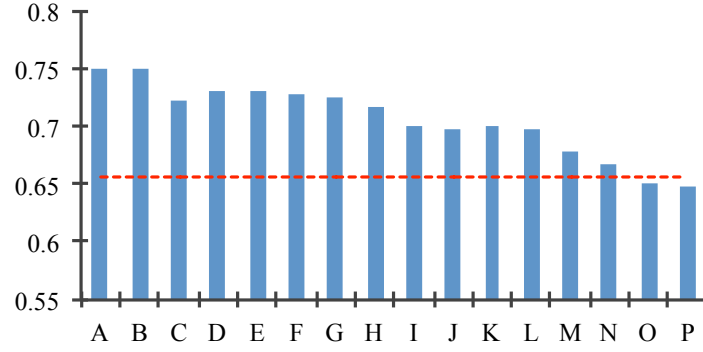


Figure 4.6: $F1$ - results of feature combination with early fusion. The red line marks the result of the best single feature. The corresponding feature combinations are denoted in Table 4.2.

Table 4.2: Feature Combinations used for early and late fusion.

ID	Features
A	LBP_Hist_u2_n16_r3, GLCMStat32, RGBHist32, EdgeHist, DSIFT_gauss
B	LBP_Hist_u2_n16_r3, GLCMStat16Order3, RGBHist32, EdgeHist, DSIFT_gauss
C	LBP_Hist_u2_n16_r3, GLCM16Order3Mean10, RGBHist32, EdgeHist, DSIFT_gauss
D	GLCM16Order3Mean10, RGBHist32, EdgeHist, DSIFT_gauss
E	GLCMStat32, RGBHist32, EdgeHist, DSIFT_gauss
F	GLCMStat16, RGBHist32, EdgeHist, DSIFT_gauss
G	GLCM16Order3Mean10, RGBHist32, DSIFT_gauss
H	GLCM16Order3Mean10, RGBHist32, EdgeHist
I	GLCMStat32, RGBHist32, DSIFT_gauss
J	GLCMStat32, RGBHist32, EdgeHist
K	GLCM16Order3Mean10, RGBHist32
L	GLCM16Order3Mean10, DSIFT_gauss
M	GLCM16Order3Mean10, EdgeHist
N	GLCMStat32, RGBHist32
O	GLCMStat32, DSIFT_gauss
P	GLCMStat32, EdgeHist

4. SEGMENTATION OF PETROGLYPH IMAGES

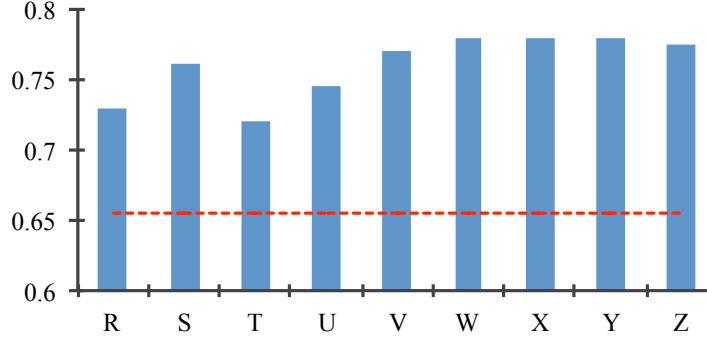


Figure 4.7: $F1$ - results of feature combination with late fusion. The red line marks the result of the best single feature. The corresponding feature combinations are denoted in Table 4.3.

4.4 Results

4.4.1 Single Features

Figure 4.5 shows selected $F1$ -results of the single feature classifiers. We observe that the LBP variants and third-order GLCM variants perform best. The RGB histogram with 32 bins is also performing well. The luminance histogram and the DSIFT variants perform poorly. The edge histogram is performing moderately. The second order GLCM variants perform comparable to the edge histogram. The statistical measures perform better than the raw GLCMs. We observe that a higher number of grey values leads to better results. We make the opposite observation in the case of the third order GLCM variants. The raw GLCMs, that have been reduced in dimensionality by discarding the sparsest dimensions, perform better than the statistical measures. A higher number of grey values results in worse performance. The statistical measures of the third-order GLCMs can only be calculated from the complete (i.e. prior dimensionality reduction) GLCMs. Hence, we assume, that a high number of grey values results in over-fitting, and therefore in lower quality results. The LBP variants deliver the best results. We observe that a bigger radius and a lower number of sample points improve results. The best result was obtained with the LBP histogram calculated from 8 sample points with a radius of 3. We conclude, that texture features generally outperform the other feature categories.

4.4.2 Early Fusion

Figure 4.6 contains selected results of early fusion classifiers. The corresponding feature combinations are denoted in Table 4.2. The number of features we combined decreases from A to P.

Bars A,B and C represent the results of combinations of five features. They contain well-performing features of each feature category. The only variation in the feature

Table 4.3: Fused classifiers. Please refer to Table 4.2 for features used in feature combinations.

ID	Feature Combinations
R	Five single feature classifiers trained with the features of combination A
S	The classifiers from R and the early fusion classifier A
T	Five single feature classifiers trained with the features of combination B
U	The classifiers from T and the early fusion classifier B
V	Six single feature classifiers trained with the features of combinations A and B and the early fusion classifiers from A and B
W	Seven single feature classifiers trained with the features of combinations A, B and C and the early feature classifiers from A, B and D
X	The classifiers from W and the early fusion classifiers E and F
Y	The classifiers from X and the early fusion classifiers G, H, I and J
Z	All single feature and early fusion classifiers

combinations of A, B and C is the used GLCM variant. A and B contain second and third order statistic measures, C contains a third order GLCM. We observe that A and B perform nearly similar, while C performs significantly worse. This is noteworthy, as the single feature performance is contrary. The GLCMStat features have 21 dimensions. The GLCM16Order3 has 4096 dimensions, the reduced version we use has 255 dimensions. We assume, that the performance drop that occurs with the utilization of GLCM16Order3Mean10 is a consequence of over-fitting.

Bars D-F show the results of combinations of four features. The combinations contain features of all feature categories except the LBPs. Bar D contains the feature combination of bar C except the LBP. We observe that D is better than C, despite the fact that it contains less information. This is more evidence for our assumption, that C performs poorly due to over-fitting. Generally, the four feature combinations perform comparably well.

Bars G-J display the results of combinations of three features. The combinations contain a GLCM variant, an RGB histogram and either an edge histogram or a dense SIFT descriptor. We observe that the combinations with GLCM16Order3Mean10 (G and H) are generally better than the combinations that utilize GLCMStat16 (I and J).

Bars K-P show the results of combinations of two features. Bars K-M contain GLCM16Order3Mean10 and a feature of another feature category, bars N-P contain GLCMStat32 and a feature of another feature category. We observe that the combinations with GLCM16Order3Mean10 outperform the combinations with GLCMStat32.

We conclude, that early fusion by simple concatenation of feature vectors significantly improves the results in comparison to single feature classifiers (see Figure 4.6). The more features we combine, the better the results are. The best results occur at the combination of features of each feature category. Over-fitting as a result of high feature dimensionality limits the number of features to be combined (see classifier C in Figure 4.6).

4. SEGMENTATION OF PETROGLYPH IMAGES

4.4.3 Late Fusion

Figure 4.7 contains selected results of late fusion classifiers. Bar R represents the late fusion result of the classifiers trained with the features of early fusion run A. We observe that in this case early fusion performs better than late fusion. Bar S shows the results of the classifiers from R and the early fusion classifier A. We observe that the result of S is slightly better than the result of A (see Figure 4.6). This is noteworthy, as it contains the same features as A, i.e. the information primarily extracted from the image is similar. Bars T and U represent the same idea as R and S, but with the features of early fusion run B instead of A. We observe that also in this case the combination of the single feature classifiers with the early fusion classifier of the same features performs better than the combination of the single feature classifiers only. In contrast to S, U performs slightly worse than the corresponding early fusion run B.

Bar V represents the late fusion result of six single feature classifiers trained with all features in A and B combined with the early fusion classifiers A and B. We observe that the combination of more features improves results.

Bars W, X, Y and Z extend the number of classifiers used for late fusion further. Bar W includes a classifier for each single feature in early fusion runs A-D as well as the three best performing early fusion classifiers A, B and D.¹ We observe, that late fusion result W significantly outperforms the best early fusion run. Bar X includes all the classifiers from W and adds early fusion classifiers E and F. We observe that X performs slightly better than W. We assume that the early fusion classifiers D, E and F contain complementary information, although they perform almost similar. Bar Y includes all classifiers from X and adds the early fusion classifiers G, H, I and J. We observe that Y performs slightly worse than X. Bar Z represents the result of late fusion including all single feature classifiers from X as well as all early fusion classifiers A-P (except C). We observe that the result is significantly better than any early fusion result, but slightly poorer than X. We assume that the poor performing early fusion classifiers I-P do not contain any information to improve the result. Nevertheless, we observe that late fusion efficiently profits from the information in the different classifiers. The results of the best classifier combination X is visualized in Figure 4.11.

We conclude, that late fusion significantly improves the results compared to early fusion. We assume, that the combination of labels in contrast to early fusion, where feature vectors are combined, allows the training of more precise classifiers, and consequently preserves the complementary information of the different features more efficiently. Above all, late fusion can be utilized to combine as many classifiers as needed, because - in contrast to early fusion - there is no danger of over-fitting.

4.4.4 Late Fusion with Thresholded Majority Voting

Figure 4.8 shows the visualization of late fusion with thresholded majority voting. We observe, that we achieve smooth segmentation results by considering also pixels with lower confidence.

¹We skip early fusion classifier C due to the assumption that it is over-fitted.



Figure 4.8: Two details of our test material with a visualization of late fusion run Y with thresholded majority voting. The grey value of a pixel corresponds to the classification confidence of a pixel. Darker means more confident. We observe that some blurring happens, e.g. at the boundaries of the shield in the upper right. We assume that the blurring is due to the size of the local window we use for feature extraction.

4. SEGMENTATION OF PETROGLYPH IMAGES

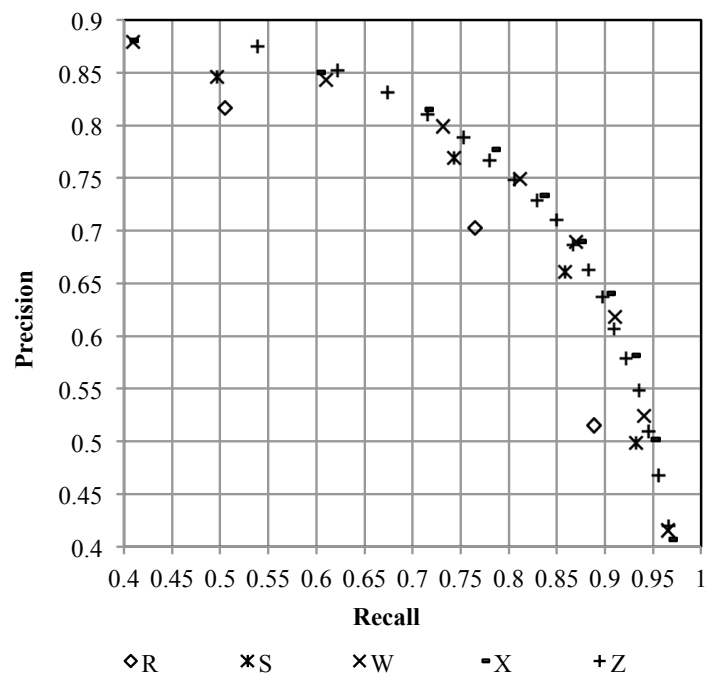


Figure 4.9: Precision and recall for late fusion with thresholded majority voting. The threshold ranges from two to the number of fused classifiers minus one. The corresponding feature combinations are denoted in Table 4.3.

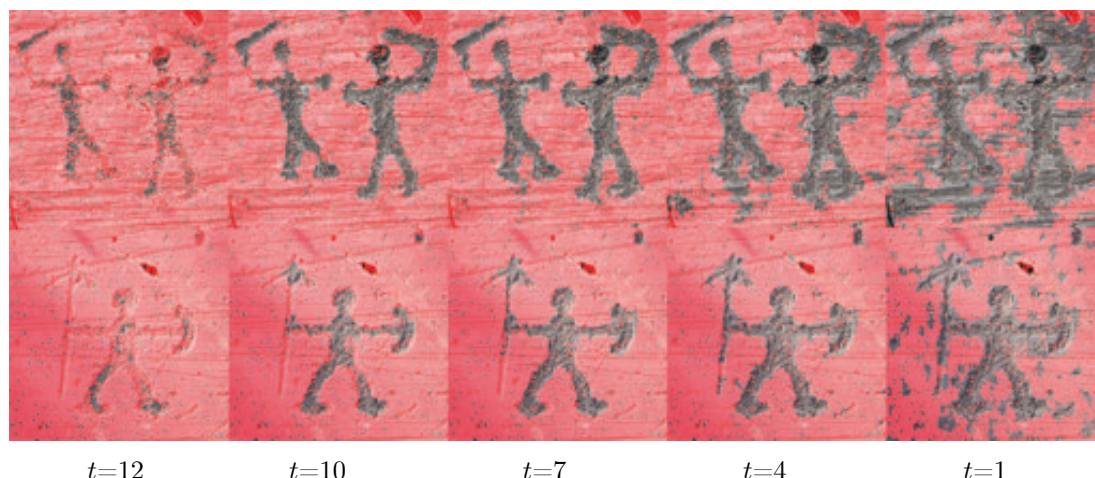


Figure 4.10: Visualization of two details of our test material classified with different thresholds used in late fusion run Y. Red dots are classified as background. Each column shows the result of a specific threshold t . The most left column corresponds to $t=12$, i.e. all fused classifiers must agree. The most right column corresponds to $t=1$, i.e. at least one classifier must classify a pixel as foreground pixel. We observe that the ideal threshold for segmentation varies.

Figure 4.9 shows the precision-recall curves of the five late fusion runs R,S,W,X and Z with different thresholds. We observe, that a change of the majority voting threshold allows us to vary our results along the precision-recall curve. This can be useful for different applications, where a balanced precision and recall are not always preferable. Moreover we observe, that with certain classifier combinations, we can move the precision-recall curve to the upper right, i.e. our good late fusion results are not caused by an optimization of the $F1$ -value.

4.4.5 Interactive Late Fusion

Figure 4.10 shows the visualization of two details of our test material overlaid with the classification results of different late fusion thresholds.

The adjustment of the thresholds allows the user to interactively increase the subjective quality of the result. Technically, the varying threshold corresponds to a varying position of the result along the precision-recall curve (see Figure 4.9). Furthermore, we observe that different parts of our test material demand different thresholds for ideal segmentation. The ideal threshold for the detail in the top row is 10, while the ideal threshold for the detail in the bottom row is 4. We assume, that this is caused by different contrast in our test material.

We conclude, that the possibility to vary the threshold interactively improves the subjective quality of the results.

4.5 Conclusion

We presented an experimental study on petroglyph segmentation by pixel classification. We showed, that with appropriate features, feature combination and fusion strategies good segmentation results are achievable. The most important lessons learned are:

- Texture features generally outperform features of other categories. The best performing single features are third order grey level co-occurrence matrices and local binary patterns.
- Early fusion significantly improves results compared to single features. The more features we fuse the better. High dimensionality of feature vectors and subsequent over-fitting limit the number of single features to be fused.
- Late fusion by fusion of classifier output labels clearly outperforms early fusion. Complementary feature information is preserved better, and there is less danger of over-fitting.

Varying the number of votes needed for the majority voting used in late fusion allows to determine a confidence for each classified pixel as well as to change the position of the classification result on the recall-precision curve. We assume, that this variable threshold improves the use of the results.

The results are good, even better when the classifier threshold is chosen manually. What we cannot solve is the dependence on the lighting during photographic acquisition of images. The images we used for our experiments are well-lit because we awaited the appropriate time of the day for capturing. We suppose that pixel classification based on 2D RGB images performs worse with different illumination. To be more independent of lighting we approach the segmentation task on 3D point clouds in the following chapter.



Figure 4.11: Our test material overlaid with the best classification result (late fusion run X).

4. SEGMENTATION OF PETROGLYPH IMAGES

Segmentation of Petroglyph 3D Point Clouds



Figure 5.1: A petroglyph scene with overlaid segmentation ground truth.

IN this chapter we address the segmentation of petroglyphs based on 3D scans. We assume that the distinctive description of the micro-topography of a surface is easier without the influence of lighting which poses a major challenge for the image-based segmentation described in the previous chapter.

The appearance of a 3D surface can be seen as the composition of *visual appearance* (visual texture, image texture) and *tactile appearance* (surface texture) [TJ98, BMdA12]. Visual appearance refers to color variations, brightness and reflectivity. Surface texture refers to the geometric micro-structure of a surface in terms of roughness, waviness, lay, coarseness, smoothness, polish, burnish and bumpiness [BJ03, BMdA12]. It is defined as the repetitive random deviation from the “ideal” surface. This deviation forms the three-dimensional *topography* of a surface [ASM96]. Compared to the visual appearance of a surface, the surface topography is invariant to lighting and viewing conditions and thus a robust basis for the description of a 3D surface. Research has mainly focused on the description and classification of surfaces in terms of their visually apparent texture, fostered by the broad availability of image texture datasets [VZ05, ZMLS07]. The analysis of surface texture directly from 3D data has been neglected so far due to the absence of suitable high-resolution 3D data.

Our petroglyph scans are highly detailed with a resolution of 0.1mm (see Figure 5.1

5. SEGMENTATION OF PETROGLYPH 3D POINT CLOUDS

for an orthophoto and Appendix A for details on the 3D dataset). Hence, they pose an opportunity to approach the analysis of surface texture based on 3D data. We follow two different strategies: In Section 5.1 we use the high resolution point clouds directly and employ 3D descriptors. In Section 5.2 we transfer the problem to image space by using a depth map of the surface. The depth map contains all necessary information due to the fact that the rock surfaces containing petroglyphs usually do not contain self-occlusions. In image space we use well-established 2D image features to distinguish between natural rock surface and pecked rock surface.

5.1 3D Descriptor-based Point Cloud Segmentation

5.1.1 Related Work

3D reconstructions are usually represented as point clouds, meshes and range images. For the representation of surface topography, descriptors are required that capture the local geometry around a given point. A large number of local 3D descriptors has been developed in recent years. Johnson and Hebert propose *spin images* as local descriptors for the dense description of meshes for object recognition in 3D scenes [JH99]. Darom and Keller introduce a scale detection scheme for mesh points and propose scale-invariant spin images [DK12]. Furthermore they extract SIFT features from local depth images of a points' neighborhood to model the surface topography around it. Zaharescu et al. introduce a local surface descriptor for meshes (MeshHOG) that can be computed from an arbitrary scalar function (e.g. curvature) defined over the surface [ZBVH09]. Steder et al. propose the *normal-aligned radial feature* (NARF) which captures local depth variations in range data [SRKB11]. Frome et al. extend the well-known 2D *shape context* descriptor [BMP02] to 3D point clouds (3DSC) and show that it outperforms spin images [FHK⁺04]. Rusu et al. propose two local 3D point cloud descriptors, namely *persistent point feature histogram* (PFH) [RMBB08] and an accelerated version *fast PFH* (FPFH) [RBB09]. Both build upon the relations between *surflets*, i.e. the combination of a surface point and its normal [WHH03]. Tombari et al. propose a 3D descriptor (SHOT) based on the point normals that is defined in a robust local reference frame [TSDS10].

5.1.2 Approach

We extract 3D descriptors based on the neighborhood of each point to describe the surface properties of this point. Subsequently, we use the descriptors and the ground truth from Subsection 2.6.1 in a machine learning setup to develop and train a model which is then used to classify points represented by their neighborhood.

The descriptors introduced in Section 5.1.1 are all suitable for the description of attributes related to surface topography although some of them were originally developed for different purposes. Out of these, we use four state-of-the-art local descriptors that exclusively describe the geometry of the surface without considering color. We select

5.1 3D Descriptor-based Point Cloud Segmentation

the descriptors following three criteria: (i) Is the descriptor local? (ii) Is the descriptor working on the point cloud only, without rendering a depth map of the considered region internally? (iii) Is it rotation invariant and robust to clutter and noise? All four descriptors meet the selection criteria. In the subsequent paragraphs we describe the four selected descriptors.

Persistent point feature histograms - PFH: The persistent point feature histograms (PFH) have been proposed by Rusu et al. [RMBB08]. PFH is a local descriptor for the geometry of the neighborhood of a given point. For a given neighborhood around a point, PFH encodes pairwise geometric relations between all surflets (a concept previously introduced by Wahl et al. [WHH03]) in the neighborhood. A surflet is defined as the combination of the position vector and the normal vector of a point. The original implementation uses angular differences between the normals based on a Darboux frame and the Euclidean distance between the points, which results in a description of each point pair that consists of four values (see Figure 5.2). In the current implementation used by us¹, the distance between the points is discarded, because it has shown to be undiscriminative. The three angles which describe the geometric relation of two points (i.e. their surface normals) are placed in a histogram which utilizes five divisions for each of the dimensions. This results in a feature vector with $5^3 = 125$ dimensions. The PFH is invariant to rotation, translation and (when discarding the Euclidean distance between the points) scale. As the comparison values are stored in a histogram, it has a certain robustness to noise. Furthermore, the authors claim, that the PFH has a robustness to the acquired scale (i.e. 3D resolution) of a scene. That might be true for the kitchen scene used for evaluation in the paper, as it does not matter whether 10 or 100 points are sampled from a flat surface in the scene and subsequently populate the histogram as described above. However, in our case, where finest detail is necessary to model the difference between pecked surface and natural surface, we do not expect the PFH to be robust to changes in 3D resolution of the acquired surfaces.

Fast Persistent point feature histograms - FPFH: The FPFH descriptor is a simplified version of the PFH [RBB09]. The simplification aims at faster computation of the descriptor. While the PFH considers the pairwise geometric differences between all points in the neighborhood of the point for which the descriptor is estimated (the query point), the FPFH descriptor in a first step only considers the pairwise differences between the keypoint and each neighbor. This descriptor is called SPFH (Simplified Point Feature Histogram). This reduces computational complexity from $O(nk^2)$ to $O(nk)$, where n is the number of points for which the descriptor is calculated, and k the number of considered neighbors. The loss (compared to PFH) of the information of pairwise relations between the neighboring points is partly compensated by addition of the keypoint's SPFH with the weighted SPFHs of its neighbors. The weight is the Euclidean distance between the query point and the neighbor. The three angular values are used in the histogram, which has eleven bins. In contrast to the PFH, the features are de-correlated, which results in a histogram of $3 \cdot 11 = 33$ feature dimensions.

Signature of Histograms of Orientations - SHOT: SHOT captures topologic infor-

¹See http://pointclouds.org/documentation/tutorials/pfh_estimation.php

5. SEGMENTATION OF PETROGLYPH 3D POINT CLOUDS

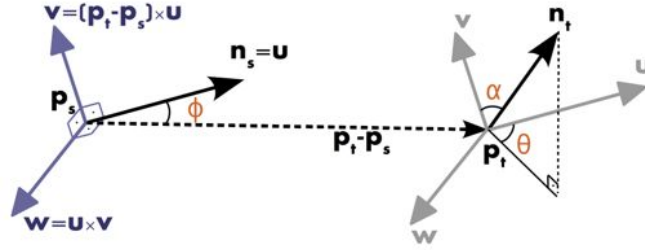


Figure 5.2: Angular features computed for a pair of points for the PFH descriptor. Vectors u , v and w form the fixed coordinate frame. α , ϕ and θ denote the calculated angular features between the normals n_s and n_t .¹

mation around a 3D point cloud by using a spherical support structure and a local reference frame [TSDS10]. The sphere around the key point is divided in 32 equally sized volumes. For each volume, a local 11 bin histogram is computed from the cosine of the angles between the key point normal and the normal of each point in the current volume. Finally, all histograms are concatenated resulting in a $32 \times 11 = 352$ dimensional feature vector. In [TSDS10], SHOT is evaluated on point clouds of synthetic 3D models. It is invariant to rotation and robust to clutter and noise. Salti et al. evaluate SHOT and other descriptors like FPFH considering the influence of the reference frame and the descriptor separately [STDS14]. They show that their reference frame performs well with other descriptors.

3D Shape Context - 3DSC: Frome et al. propose the 3D Shape Context descriptor as a local point descriptor enabling object recognition in range data [FHK⁺04]. The motivations for the descriptor are cluttered and noisy scenes acquired by range scanners. The evaluation of 3DSC is performed on synthetic 3D car models with added noise. The descriptor is a straightforward 3D generalization of the well-known Shape Context descriptor proposed by Belongie et al. [BMP02]. A spherical support volume is divided radially and around the azimuth in bins. The occurrences of points in the divisions are accumulated in a histogram. To provide rotation invariance around the azimuth, the histograms are rotated in order to have a descriptor for each division being the start division. The resulting histograms are concatenated. We use an 1980-dimensional feature vector in our experiments.

The dense computation of 3D descriptors is computationally demanding. Hence, we select a subset of our 3D data for the evaluation. We perform experiments on four scans of our 3D point cloud dataset. See Appendix A for details on the 3D dataset. We use the scans with the IDs 1, 14, 21 and 22.

Qualitative evaluation of the descriptors: The descriptors we investigate for our task have originally been designed to work on scenes like a kitchen [RMBB08] or for similarity estimation of synthetic 3D models [FHK⁺04]. Our task (surface classification) is different to these typical tasks where large surface parts of the point clouds are close

¹Image source: <http://pointclouds.org>.

5.2 Image-space-based Point Cloud Segmentation

to volumetric primitives. In our case, we have small structures in the stone (peck marks) that we want to use for the distinction of natural rock surface and pecked rock surface. Hence, our first look into the descriptors is qualitative. In this first step, we aim at investigating the discriminative capability of each descriptor without taking generalization ability into account. We train a classifier with a randomly selected training set from one image and classify the remaining feature vectors with the model obtained.

Cross-Validated Evaluation of Single Descriptors and Descriptor Combinations: We evaluate the performance of single features and concatenated combinations. Randomly selected feature vectors of point clouds 14 and 21 (see above) are the training set, the two other point clouds as well as the remaining points from clouds 14 and 21 form the test set. As we are operating on the full resolution point clouds (the small dataset consisting of four of the 26 scans has more than 12 million points) we utilize only 2.5% of point clouds 14 and 21 as training data. For validation we use repeated random sub-sampling with 10 sub-samples [Rao05].

5.1.3 Results

Figure 5.3 shows the qualitative results of the experiments. Generally, we observe, that the classifier trained on each of the local point cloud descriptors produces lots of regions that are false positives, i.e. regions that are natural rock surface but are classified as pecked areas. A closer look at the single descriptors results shows, that PFH detects most of the pecked areas but has a large amount of false positives as well. FPFH performs comparable in terms of false positives, but misses large pecked regions. SHOT and 3DSC produce noisy results over the whole 3D scan, with 3DSC showing more high frequency noise than SHOT.

The boxplots in Figure 5.4 reflect the 10 $F1$ -values obtained in the quantitative evaluation. We observe, that the Point Feature Histogram (PFH) clearly outperforms all other descriptors with an $F1$ -value around 0.45. The simplified descriptor FPFH is more than 0.1 behind with an $F1$ -value around 0.31. SHOT and 3DSC are both around an $F1$ -value of 0.25. The combination of descriptors with early fusion does not improve results at all. Combinations containing PFH (the first three combinations) perform comparable to PFH only, and the fourth combination (FPFH+SHOT) performs comparable to FPFH alone. We assume that this results from the fact that all descriptors exploit similar information, the geometric relations of points and the respective surface normals. We avoided combining 3DSC due to its large dimensionality (1980 dims).

5.2 Image-space-based Point Cloud Segmentation

In this section, we investigate the problem of describing and classifying 3D surfaces according to their topographic structure in high-resolution surface reconstructions. This task is challenging for the following reasons: First, methods such as SfM produce densely sampled reconstructions which may contain several millions of points. The dense ex-

5. SEGMENTATION OF PETROGLYPH 3D POINT CLOUDS



Figure 5.3: Qualitative evaluation of the discriminative abilities of the employed 3D point cloud geometry-based descriptors. The rows from top to bottom show the ground truth, PFH, FPFH, SHOT and 3DSC. The right hand side shows details. We observe, that each point cloud descriptor produces a large amount of false positives compared to the ground truth in row 1.

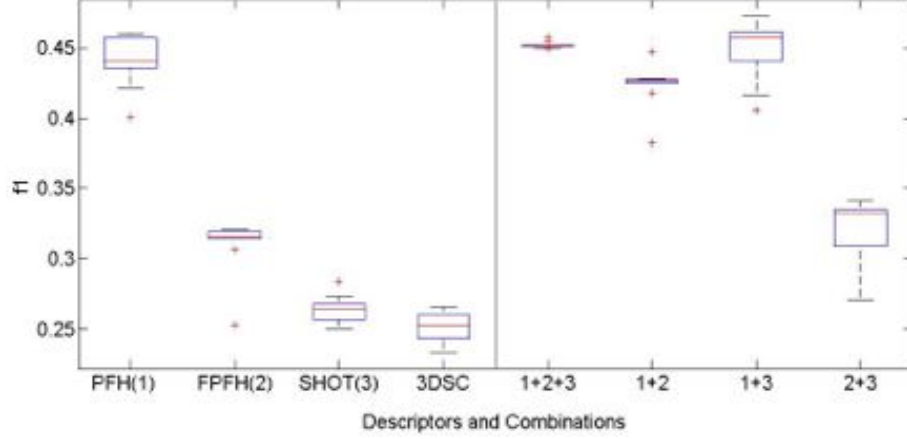


Figure 5.4: Segmentation results with different local point cloud surface descriptors. The left side shows the performance of single features, the right side shows the performance of descriptor combinations. We observe, that the combinations yield only slight improvements.

traction of a surface’s topography in 3D is a computationally demanding task that may lead to unacceptable runtimes for many applications. Second, surfaces may be shaped differently and exhibit a complex nonlinear geometry (curvature) at the global level that has to be compensated prior to the analysis of the local topography.

The dense 3D descriptor extraction for experiments of the previous section took several days to complete. Therefore it is not feasible for larger datasets. As a consequence, we transfer the problem to image space. The highly detailed scans of surface areas we use do not include self-occlusions (see Appendix A). Hence, it is possible to reduce the data from 3D to 2.5D (plane with height information, i.e. depth map) without any information loss. The depth map and a derived gradient map can form input to standard state-of-the-art 2D pixel-wise segmentation methods as the one described in Chapter 4 and can therefore build on long developed and highly sophisticated 2D features.

5.2.1 Related work

The analysis of surface topography is necessary in different application domains. Applications include surface matching [JH99], automated terrain classification for autonomous vehicles, inspection of material quality in industrial applications (an important fundament for this line of work are the “Birmingham 14” parameters for characterizing three-dimensional surface topography [DSS92, DSS93, DSS94a, DSS94b]), analysis of digital elevation models in geology (geomorphometric analysis, a recent overview can be found in [HR08]), as well as the analysis of archeological artifacts [BMdA12, Sco12]. Surface topography (sometimes referred to as morphology) may be investigated at different scales. While inspection of material quality may be performed at a millimeter scale or even below, geologic morphology analysis is usually performed at a scale of

5. SEGMENTATION OF PETROGLYPH 3D POINT CLOUDS

meters or kilometers. While the data lives in different scales, morphological parameters are in general invariant to scale. Morphological parameters have been defined at several places in literature and depend on the application domain. A broad range of parameters has been developed for geomorphometry [Lea11]. A categorization of parameters related to archeology are presented in [BMdA12]. Independent of the application domain, the parameters are comparable to a high degree. Further parameters are defined in [BJ03].

5.2.2 Descriptors

There are numerous 2D descriptors that can be utilized for this analysis. Additionally to the 2D descriptors described in Subsection 4.3.2 we use the Histograms of Oriented Gradients (HOG, [DT05]) and three simple and easy to compute features proposed for the description of enhanced deviation maps in [ZS15]: (i) global histogram shape (GHS), (ii) tiny image (TI), and (iii) spatial frequencies (SF). All features are defined for a local neighborhood. The size of the local window is 128x128px which is equivalent to 10.8x10.8mm² (see Section 2.5 for details on orthophoto generation). GHS is computed from a global histogram of all deviation values in a given neighborhood. To capture solely the shape of the histogram, we extract the first 30 discrete cosine transform (DCT) coefficients from the histogram and use them as a feature vector to describe the global value distribution in the local neighborhood. TI is simply a down-scaled 8x8 pixel version of the current neighborhood and is inspired by [KJPK02] who have shown that the raw pixel values yield surprisingly good performance for 2D texture analysis when used as features. The third feature, SF, contains the first 8x8 low frequency 2D DCT coefficients of a given image block. SF actually captures the same information as TI in a different mathematical domain (frequency domain). We employ both features to evaluate which representation suits topography classification better. While GHS represents spatial invariant information, TI and SF capture mostly spatial information. Thus both types of features are complementary which makes them good candidates for combination.

5.2.3 Approach

A depth map encodes global (the curvature of the *ideal surface* as described at the beginning of the chapter) as well as local (the high frequency *deviations* of the surface from the ideal surface) curvatures of the surface. Figures 5.5(a) and 5.5(b) show a point cloud and the corresponding depth map color coded as heat map. We observe that the global curvature of the point cloud has larger differences in values than the local depth differences resulting from the pecked figures. For the distinction of pecked and natural parts of the rock surface we are interested in high frequency local curvatures and have to be robust against the global curvature. To accomplish this we follow two strategies:

a) *Local feature extraction*: We extract features in a block around the pixel we want to classify as natural rock or pecked rock. We normalise the feature values and hence compensate for the global height of the block. But, there is one thing we cannot compensate with local feature extraction: A global inclination of the surface patch

5.2 Image-space-based Point Cloud Segmentation

which is the basis for the depth map block used. We assume, that different features may be robust against such constant gradients while others rely strongly on gradients. Hence, we employ different features and feature combinations in the experiments.

b) *Global curvature compensation and local curvature enhancement*: In order to be robust against the global curvature resulting in different inclinations of the surface patches for feature extraction we compensate the global curvature by subtraction of a Gaussian-filtered version of the depth map from itself. The local curvature is enhanced by several image processing steps. The process from the depth map to the enhanced depth map has been proposed and evaluated by Zeppelzauer and Seidl in [ZS15]. Figure 5.5 illustrates the topography extraction and enhancement on a surface from our dataset. The point cloud viewed from projection direction together with its depth map is shown in Figures 5.5(a) and 5.5(b). The deviation map in Figure 5.5(c) compensates the global depth variations and reveals the micro-structure of the surface. The corresponding enhanced deviation map in Figure 5.5(d) improves basic topographic patterns, such as valleys in this example and yields an expressive representation of the surface topography.

5.2.4 Experimental setup

The experiments aim at comparing the full 3D point-based approach that relies on different local 3D descriptors extracted directly from the point cloud described in Section 5.1 with a pixel classification task on depth maps and derivatives thereof. First, we compare the full 3D approach with different 2D input representations (e.g. depth maps, color images, enhanced depth map) and several state-of-the-art 2D features extracted from them. Additionally, we employ the three features described in Subsection 5.2.1 (GHS, TI, SF) extracted from the enhanced deviation map.

The employed dataset contains 26 surface reconstructions with a total number of 115.3M points (see Appendix A for details on the 3D dataset). The number of points per surface ranges from 810k to 10.6M. For each surface a precise ground truth has been generated by domain experts that labels all engravings on the surface (see Subsection 2.6.1). The dataset contains two classes of surface topographies: class 1 represents engraved areas and class 2 represents the natural rock surface. Class priors are imbalanced. Class 1 represents 20.9% of the data and is thus underrepresented.

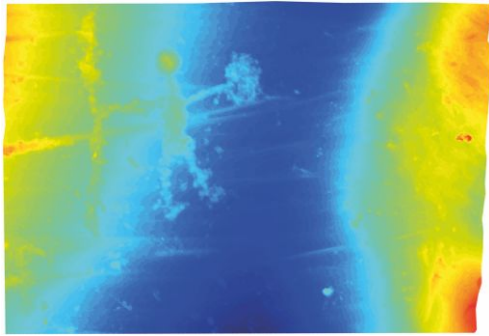
In a first step, we perform all experiments on the four scans used in the full 3D evaluation in Section 5.1. Then we employ the most promising features on the full dataset.

For the representation of the input point cloud, we investigate four different visual representations: (i) A color image of the point cloud obtained by orthographic projection of the point colors. The point colors are obtained by shading the point cloud with oblique illumination. This makes topographic patterns better visible in the resulting texture of the image. (ii) The depth map obtained by orthographic projection. (iii) The gradient map which is the local gradient of the depth map. The gradient emphasizes local depth variations and thus captures structures related to surface topography. At the same time it neglects global surface variations. It has been successfully employed

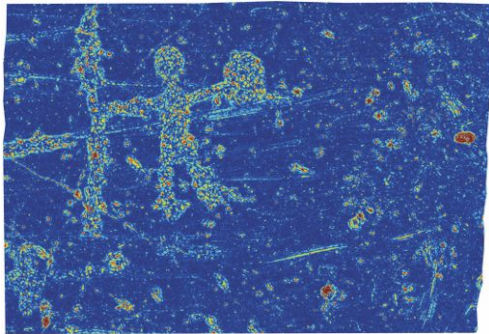
5. SEGMENTATION OF PETROGLYPH 3D POINT CLOUDS



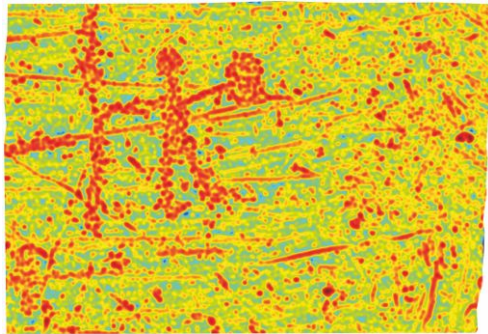
(a) Visual appearance.



(b) Depth map.



(c) Deviation map.



(d) Enhanced deviation map.

Figure 5.5: A 3D surface reconstruction of a rock surface in which the shape of a human figure has been engraved. (a) The colored point cloud; (b) The projected depth map; (c) The deviation map; (d) The enhanced deviation map accentuates well the different surface topographies of the engraving and the surrounding rock surface. Please refer to [ZS15] for details on the process from depth image to enhanced deviation map.

5.2 Image-space-based Point Cloud Segmentation

for the segmentation of fine structures in 3D scenes [WFFD11]. (iv) The enhanced deviation map from [ZS15].

From all four representations, we compute the following features: MPEG-7 Edge Histograms (EH) [II02], two variants of Local Binary Patterns: uniform LBP (uLBP) and rotation invariant LBP (riLBP) [OPH96, OPM02], Scale-Invariant Feature Transform (SIFT) descriptors [Low04], Histograms of Oriented Gradients (HOG) [DT05] and Gray-Level Co-occurrence Matrices (GLCM) with derived parameters specified in [HSD73, ST99, Cla02]. Additionally, we extract the features proposed in Subsection 5.2.1 (GHS, TI, and SF) as well as two combinations (GHS+TI and GHS+SF) which are formed by concatenating the individual feature vectors. All features are extracted for local windows of 128x128 pixels ($10.8 \times 10.8 \text{mm}^2$) with a step size of 16 pixels (1.35mm).

For classification we employ SVMs (with linear and RBF kernel) [CV95], and random undersampling boosting (RUSBoost) which is a variant of AdaBoost [FS⁺96] that is especially designed for imbalanced class distributions [SKVHN08, SKVHN10].

We split the entire dataset into two sets: a set of training surfaces and a set of test surfaces. Training sets are randomly selected from the training surfaces. We apply 5-fold cross-validation and grid search on the training sets to determine the classifiers' parameters (C for linear SVM, C and γ for SVM with RBF kernel, and the number of learners for RUSBoost). The classifiers with optimized parameters on the cross-validated training set are then applied to the previously unseen test surfaces. We compute recall and precision for both classes and report the $F1$ -score of the *underrepresented* class for each experiment. As class 2 covers nearly 80% of the surfaces' area, the $F1$ -score of the underrepresented class (class 1) is most expressive to assess the overall classification performance. Each experiment is repeated 10 times with different randomly selected training sets from the training surfaces to obtain a distribution of $F1$ -scores for each experiment. We report the median $F1$ -value, as well as the interquartile range (iqr) of the distribution. To judge the performance difference between two experiments, we apply Fisher's randomization test [Fis35, SAC07] as well as Student's paired t-test [BHH78] with a significance level of 0.05, i.e. in all experiments a p -value ≤ 0.05 is considered statistically significant.

5.2.5 Results

We perform a systematic evaluation of all employed 2D and 3D representations and features.

5.2.5.1 Small dataset

As the number of experiments is large, we first use the four heterogeneous surfaces used in Section 5.1 with the full 3D approach. Training data is randomly selected from the first two surfaces. The remaining two surfaces form the test set. The results in Table 5.1 (calculated from the results in Subsection 5.1.3) show that PFH is the most suitable feature for our task. PFH outperforms the other descriptors significantly

5. SEGMENTATION OF PETROGLYPH 3D POINT CLOUDS

Table 5.1: Results of the local 3D descriptors with linear SVM. Numbers provide the median $F1$ as well as the inter quartile range (iqr) in brackets. The PFH descriptor best captures the topographic surface characteristics, however, at the cost of considerable computation time.

Feature > Repres. ∨	PFH $F1$ -median (iqr)	FPFH $F1$ -median (iqr)	SHOT $F1$ -median (iqr)	3DSC $F1$ -median (iqr)
3D point c.	0.441 (0.020)	0.316 (0.005)	0.264 (0.011)	0.253 (0.016)

(p -value: 0.016). PFH represents a global statistic of the normal orientations by considering *all* points in a local neighborhood in a pair-wise manner [RC11]. FPFH is a simplified (and much faster) variant of PFH which operates on a smaller subset of pairwise comparisons. This, however leads to a coarser representation of the surface topography which negatively affects performance. The remaining features (SHOT and 3DSC) show poor performance. In contrast to PFH, SHOT and 3DSC encode spatial information from the local neighborhood which shows to be less-suited for topography description.

Then, we evaluate topography classification in 2D on the four representations used in Section 5.1 to compare the results with the full 3D results. Table 5.2 shows the results of all 2D features with all input representations with a linear SVM. Comparing the results with the 3D results in Table 5.1 we observe, that the best feature in the 3D approach performs comparably well with the color image-based segmentation but is outperformed by all representations of the point cloud considering depth. Hence, we do not consider the full 3D approach in the remaining section but concentrate on the image space methods.

In Table 5.2 we observe significant performance differences between the four input representations. The classification based on the color image yields the weakest results. The best feature is SIFT with a median $F1$ of 0.495 followed by SF and both LBP variants. The improvement of SIFT over SF is however not significant (p -value=0.56). Topography classification on the depth map yields an improvement of 5.4% for the best feature (SF with a median $F1$ of 0.549). This performance differences between depth map and color image are, however, not significant. This applies also to the gradient map and the depth map. The best feature on the gradient map (riLBP, median $F1$ =0.549), however, significantly outperforms all other features on the color image as well as all features on the gradient map (in both cases with a p -value<0.001).

Experiments on the enhanced deviation maps (EDM) show that the first 6 features in Table 5.2 cannot benefit from the enhanced representation. GLCM benefits most from the enhanced representation and is significantly better on EDM than on all other representations (p -value=0.023). The three proposed features GHS, TI, and SF yield similar performance as the top performing features on the other representations. GHS yields the best performance in terms of median $F1$ so far (0.589). This is, however, not significant, as the feature shows a strong variation in performance (iqr of 0.447). A

5.2 Image-space-based Point Cloud Segmentation

Table 5.2: Results of the systematic evaluation of all input representations and features with linear SVM on the small dataset (the four point clouds on which the full 3D approach has been evaluated in Section 5.1). Numbers provide the median $F1$ as well as the inter quartile range (iqr) in brackets. ¹ denotes that the result is significantly better than all other results for this feature, i.e. best result of a row. ² denotes that the result is significantly better than all other results for this representation, i.e. best result of a column. GHS+SF applied to the EDM significantly outperform all other features in the experiments.

Repres. > Feature \vee	Color image $F1$ - median (iqr)	Depth map $F1$ - median (iqr)	Gradient map $F1$ - median (iqr)	Enh. dev. map $F1$ - median (iqr)
EH	0.301 (0.002)	0.271 (0.000)	0.364 (0.001) ¹	0.286 (0.002)
riLBP	0.462 (0.005)	0.491 (0.006)	0.549 (0.005) ^{1,2}	0.290 (0.001)
uLBP	0.454 (0.014)	0.413 (0.073)	0.304 (0.319)	0.019 (0.022)
SIFT	0.495 (0.041) ¹	0.269 (0.003)	0.345 (0.062)	0.424 (0.018)
HOG	0.354 (0.014)	0.451 (0.009) ¹	0.414 (0.007)	0.351 (0.002)
GLCM	0.389 (0.134)	0.271 (0.000)	0.316 (0.044)	0.501 (0.184) ¹
GHS	0.192 (0.169)	0.214 (0.077)	0.321 (0.180)	0.589 (0.447)
TI	0.399 (0.042)	0.476 (0.052)	0.273 (0.001)	0.522 (0.001)
SF	0.481 (0.033)	0.549 (0.251)	0.279 (0.005)	0.521 (0.001)
GHS+TI	0.212 (0.158)	0.205 (0.083)	0.321 (0.180)	0.682 (0.104) ¹
GHS+SF	0.422 (0.034)	0.333 (0.002)	0.356 (0.037)	0.745 (0.024) ^{1,2}

5. SEGMENTATION OF PETROGLYPH 3D POINT CLOUDS

strong gain in performance is obtained when the features are combined. Both combinations GHS+TI and GHS+SF significantly outperform all other features on the EDM as well as all other features on the remaining representations (all p -values <0.047). The peak performance is obtained by GHS+SF with a median $F1$ of 0.745.

The strongly varying performance of the different features has several reasons. The enhanced deviation map is a comparably smooth representation without sharp edges (due to Gaussian filtering). Thus, features that rely on gradient orientations, such as HOG, SIFT, EH can hardly benefit from the novel representation. The GLCM that takes the absolute values of the map as input (instead of differences or gradients) benefits most from the enhanced deviation map compared to the other representations. We observe the same behavior for GHS and TI which both rely on absolute values. SF performs similar to TI because they actually represent equivalent information. The high performance of the combined features (GHS+TI and GHS+SF) has two reasons: First, both features perform well when applied separately; Second, both features represent complementary information. GHS represents the shape of the value distribution and thus is a global and rotation invariant descriptor of an image block. TI and SF, in contrast, represent the spatial distribution of values and thus add a lot of discriminatory information to GHS.

We repeat all experiments with an SVM with RBF kernel. The nonlinear SVM achieves lower training errors than the linear SVM which shows that RBF is able to better fit to the training data. During testing, however, no significant performance gain is achieved by the RBF kernel. The best feature is again GHS+SF on EDM with a median $F1$ of 0.737. We assume that overfitting prevents a further increase of performance.

We perform additional experiments with RUSBoost and summarize results in Table 5.3. RUSBoost improves performance in 40 of the 44 experiments shown in Table 5.3. RUSBoost better models the data and yields a higher generalization ability than the SVM in our experiments. The weakest representation shows to be color. For the depth and gradient map, however, a new peak performance is obtained. HOG on the depth map improves by 22.9% to 0.680 and riLBP yields a median $F1$ of 0.602 (+5.3%). The best overall results are again obtained for the enhanced deviation maps and the proposed features. RUSBoost significantly improves the results for the individual features (GHS, TI, and SF) over SVM with p -values <0.001 . Finally, RUSBoost trained on GHS+TI on EDM significantly outperforms all other experiments performed so far with a median $F1$ of 0.768 (p -value <0.001). Apart from the improved overall performance which indicates a better generalization ability, RUSBoost yields a much lower spread in performance (maximum iqr 0.071) compared to SVM (maximum iqr 0.447). This demonstrates a high robustness of RUSBoost to different training data selections.

5.2.5.2 Full dataset

We finally evaluate the proposed method for the full dataset containing all 26 surface reconstructions with a total of 115.3M points. We employ the first 10 surfaces for training and the remaining 16 for testing and pick training sets again randomly from the

Table 5.3: Results of the systematic evaluation with RUSBoost (same schema as Table 5.2).

Repres. > Feature \vee	Color image $F1$ - median (iqr)	Depth map $F1$ - median (iqr)	Gradient map $F1$ - median (iqr)	Enh. dev. map $F1$ - median (iqr)
EH	0.345 (0.004) ¹	0.000 (0.000)	0.323 (0.013)	0.334 (0.004)
riLBP	0.517 (0.007)	0.653 (0.004) ¹	0.602 (0.006) ²	0.331 (0.003)
uLBP	0.400 (0.016)	0.550 (0.028) ¹	0.495 (0.005)	0.367 (0.008)
SIFT	0.386 (0.010)	0.243 (0.002)	0.431 (0.022)	0.472 (0.005) ¹
HOG	0.442 (0.007)	0.680 (0.006) ^{1,2}	0.383 (0.005)	0.368 (0.005)
GLCM	0.397 (0.021)	0.000 (0.000)	0.391 (0.012)	0.656 (0.003) ¹
GHS	0.398 (0.007)	0.517 (0.007)	0.358 (0.064)	0.690 (0.004) ¹
TI	0.513 (0.006)	0.569 (0.011)	0.391 (0.048)	0.729 (0.003) ¹
SF	0.488 (0.004)	0.534 (0.008)	0.345 (0.005)	0.689 (0.002) ¹
GHS+TI	0.434 (0.003)	0.598 (0.007)	0.454 (0.071)	0.768 (0.004) ^{1,2}
GHS+SF	0.411 (0.008)	0.521 (0.011)	0.455 (0.010)	0.744 (0.003) ¹

training surfaces. We apply RUSBoost with 70 learners on the features that performed best so far (riLBP, HOG, GLCM, GHS, TI, SF, GHS+TI, and GHS+SF). Figure 5.6 shows boxplots of the results for all features on all representations. We observe that the overall results decrease which is not surprising, as the entire dataset contains 15 times more data than the subset evaluated so far which increases the complexity of the classification task. The individual results, however, show similar behavior as in the previous experiments. Again the color image is the weakest representation and riLBP and HOG yield the best performance on the depth and gradient maps. The enhanced depth map again yields a significant improvement of the overall performance. Only the proposed features yield $F1$ -scores above 0.6. The best result is obtained by GHS+TI (median $F1=0.637$) which significantly outperforms all other features with p -values < 0.030 . The results on the entire dataset demonstrate that the enhanced deviation map outperforms the other representations and expresses well the topography of the additional surfaces in the full dataset. The proposed features are most effective in combination with the enhanced deviation maps and best capture the topographic information of the surfaces.

5.3 Conclusion

In this chapter we investigated the segmentation of rock art 3D scans in natural rock surface and pecked rock surface based on the micro-topography of the surface. We evaluated the performance of state-of-the-art local 3D surface descriptors in comparison with image-space representations of the point clouds on our task of rock art segmentation. A comparison of the results obtained with 3D point cloud descriptors with the results on orthographic images with and without the utilization of depth on the same

5. SEGMENTATION OF PETROGLYPH 3D POINT CLOUDS

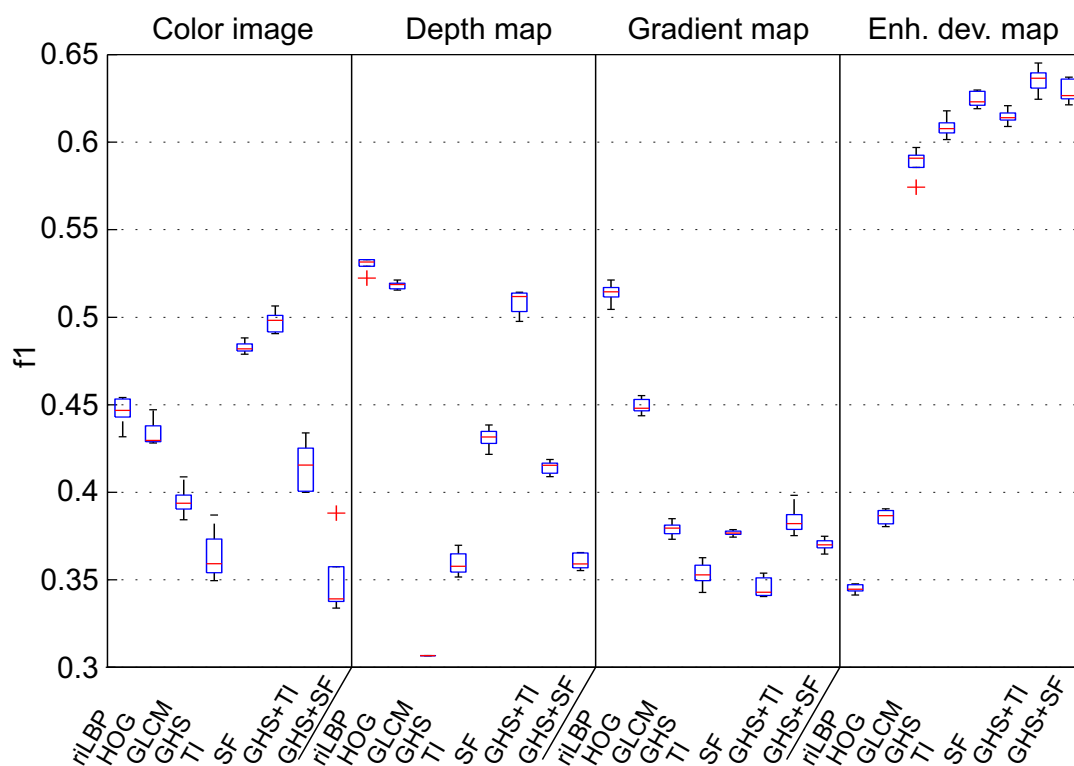


Figure 5.6: Results of the systematic evaluation with RUSBoost on the complete dataset with 26 3D scans. The enhanced deviation map yields again a significant improvement of performance.

dataset shows, that the best performing point cloud descriptor PFH is comparable with the segmentation based on RGB orthographic images and is outperformed by the best results on the non-enhanced depth image, the depth map (DM). The image space results on the gradient map (GM) and the novel enhanced depth map (EDM) are clearly better than the results achieved with PFH.

Our experiments reveal two key findings: First, surface topography classification on the point clouds directly using local 3D descriptors performed worse than expected which suggests that existing descriptors do not capture surface topography well. Second, using image space methods, an expressive representation is more important than a sophisticated feature for topography classification. In our experiments the simplest features achieved the best results on the enhanced maps.

We conclude that in terms of micro-topographic surface description, 3D is *not* here¹. A significant amount of future work has to be done in the development of 3D descriptors. Especially more efficiently extractable features are needed for a feasible employment of 3D descriptors.

¹We refer to Rusu and Cousins who titled their paper on the point cloud library "3D is here ..." [RC11].

5. SEGMENTATION OF PETROGLYPH 3D POINT CLOUDS

Similarity of Pecking Styles



Figure 6.1: A part of a petroglyph depicting a human with a superimposed arrow containing different pecking styles.

IN Chapters 4 and 5 we analysed the rock surface aiming at a distinction of pecked and non-pecked areas. In this chapter, we use the segmentation results to look at the pecked regions aiming at a distinction of different *pecking styles*, that is different depth, shape, size and spatial distribution of peck marks (see Figure 6.1). What are the reasons for different peck marks? The two main influences on peck marks are *time* and the *manufacturing process*. First, depending on the type of the rock, the time passed since the manufacturing of the pecking and the environmental conditions, the rock surface shows different types and amounts of *abrasion*. This abrasion alters the peck marks, it makes them less deep and smooths away sharp edges. Hence, the abrasion can be an indicator for the *time* passed since the manufacturing of the pecking, given that data about the past environmental conditions exists. Second, the process of creation of rock art can cause many different peck marks. Influences on the peck marks are *a) the tools used*: Stone or metal, type of stone or metal, directly pecked or with a second

6. SIMILARITY OF PECKING STYLES

object used as hammer?; *b) the author him- or herself*: Right-handed or left-handed, strong or weak, sitting or standing, in a comfortable position while working, intention to make deep or shallow peck marks?

The traditional way to estimate the properties of the manufacturing process is based on experimental archaeology. The hypotheses about the tools and other properties of the peck marks under investigation are checked by performing the creation process on a comparable rock (ideally a piece of rock from the locality of the investigated peck marks) and a subsequent visual comparison of the original peck marks with the experimentally created peck marks. For this process, an automated similarity estimation of the peck marks based on high-resolution 3D scans can be supportive. However, this is only feasible with a small number of figures. The number of parameters and variations thereof creates a large possibility space of plausible author/tool/intention combinations. Hence, it is not feasible to create experimental examples for large amounts of petroglyphs, as for instance in Valcamonica where up to 300.000 figures have been hammered into the rock. Besides the estimation of tools and other properties of the manufacturing, the second archaeological motivation for the investigation of peck marks is the study of *superimposition* which can help with relative timing. Many figures are pecked over other figures, causing merged figures (see Figures 6.2 and 6.3). The determination of the figure that was pecked last, i.e. over the other figure, can be supported by investigating the pecking style. The figure that has the same pecking style as the overlapping area of the two figures is supposed to have been pecked later (see Subsection 2.3.2).

The amount of figures with superimpositions in our 3D dataset is small and we do not have any ground truth for different pecking styles. Hence, in our approach we investigate whether different pecking styles can be distinguished based on numerical descriptions. As there is no ground truth, we use unsupervised methods (clustering) on point-wisely extracted 3D surface descriptors and visualise the results by coloring the points of the point cloud.

6.1 Related Work

Zotkina et al. conducted an experimental study of percussion and tool types in the Minusinsk basin in Siberia [ZTKP14]. They replicated different peck marks with a diverse set of tools. Moitinho et al. approached a numeric analysis of pecking styles using terminology and software from geomorphology [dARBP13]. They aim to estimate the tools and gestures used for creating rock art. The approach includes experimental archaeology to replicate ancient peck marks. They scan the surfaces with 0.28mm resolution. From the scans, they use patches with 50x50mm² containing homogeneous peck marks to extract surface metrology parameters with proprietary software. They cluster the numeric descriptions of the peck marks and observe three clusters which contain surfaces with the following properties: 1) fine, 2) coarse and regular, 3) coarse and irregular. They do not observe clusters linked to the manufacturing process of the samples. In contrast to Moitinho et al. we approach peck mark similarity estimation with 3D surface descriptors which have been developed for registration and object detec-



Figure 6.2: A warrior figure containing two visually clearly distinguishable pecking styles. We observe that the large arrow from left to right has been pecked over the warrior with clearly larger peck marks. This figure serves as a good basic test case whether 3D surface descriptors can help to automatically distinguish between pecking styles.

6. SIMILARITY OF PECKING STYLES



Figure 6.3: A scene containing at least two clearly distinguishable pecking styles.

tion. To our knowledge, there is no work on automated peck mark similarity estimation comparable to ours.

6.2 Approach

We follow two approaches to investigate peck style variety based on the point cloud data. First, we densely sample the point cloud, i.e. we compute a descriptor for each point of the cloud based on its neighborhood (see Subsection 6.2.1). The second approach (see Subsection 6.2.2) computes a topology description of a point neighborhood not for all points, but only for relevant points. For this purpose, we use the information about peak positions which we determine by the detection of local maxima in the orthographic depth map. In both cases, we cluster the descriptors and visualize the results. As we have millions of descriptors in up to 125 dimensions (in the case of PFH) we use a partitioning clustering method, namely k -means. k -means allows to cluster large datasets in reasonable time. In comparison to other clustering methods the disadvantage of k -means (and other partitioning methods) is the necessary selection of a value for k , i.e. the number of clusters. Our samples contain at most 3 visually distinguishable pecking styles. We cluster all our samples with 2,3,4 and 5 clusters to allow over-segmentation and to possibly detect visually not observable clusters.

6.2.1 Dense sampling

Based on the results of Section 5.1 we employ PFH and FPFH (Fast Persistent Point Feature Histograms) point cloud descriptors. PFH performed best while FPFH provided the best trade-off between segmentation performance and computation time. Please refer to Section 5.1 for details on 3D point cloud descriptors.

We extract PFH and FPFH densely and vary the size of the neighborhood considered for descriptor computation. We cluster the resulting descriptors in 2,3,4 and 5 clusters and perform visual inspection to assess whether the descriptor allows to model peck style similarity well.

6.2.2 Topology-based sampling

The dense approach presented in the previous section has drawbacks. First, the computational complexity of dense descriptor calculation is high. Second, the descriptors are highly redundant, as typical diameters of the marching sphere we use for the description of a point neighborhood are between 4 and 8 mm whereas the point clouds we use are rasterized with a grid size of 0.1 mm.

An alternative approach is to restrict analysis to salient points that can serve as window centers for surface description. Meaningful salient points for the description of surface texture are peaks and valleys. Figure 6.4 shows a surface part with peaks marked red. Please note, that we refer to *peaks* as maxima in the depth map values. The peaks are actually valleys if we look at the 3D surface as the detector finds peaks

6. SIMILARITY OF PECKING STYLES

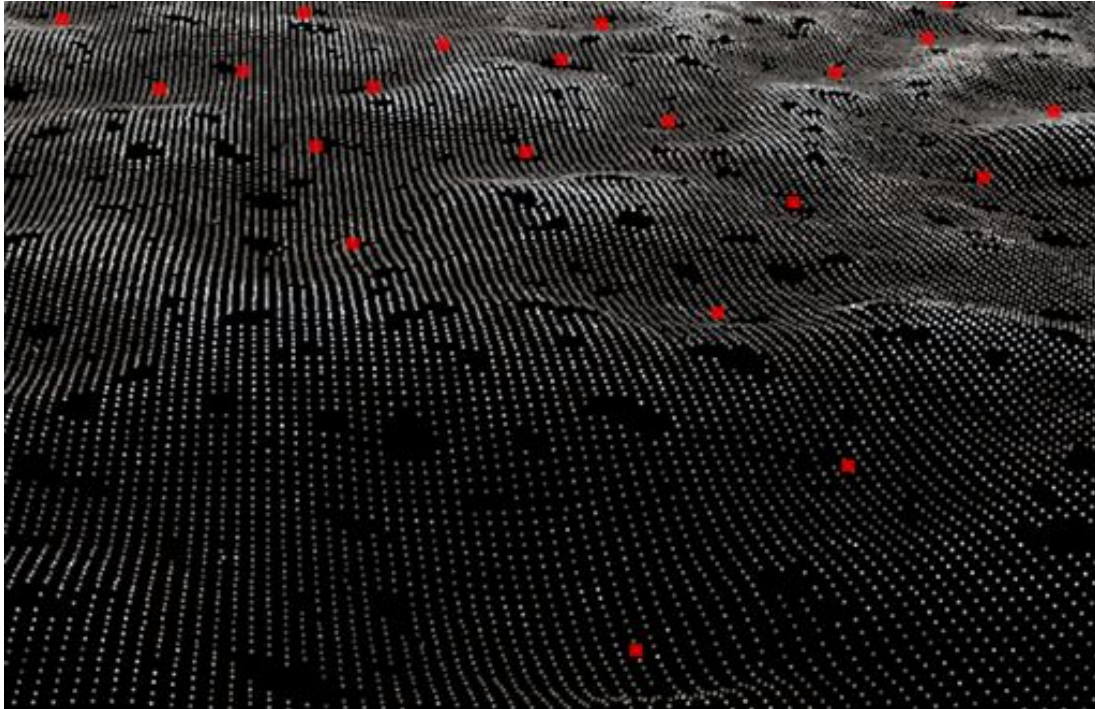


Figure 6.4: Detail of a point cloud with points at peak positions marked red. Please note, that we refer to *peaks* with respect to the detection method. The peaks are actually valleys if we look at the 3D surface as the detector finds peaks at local maxima in the depth map.

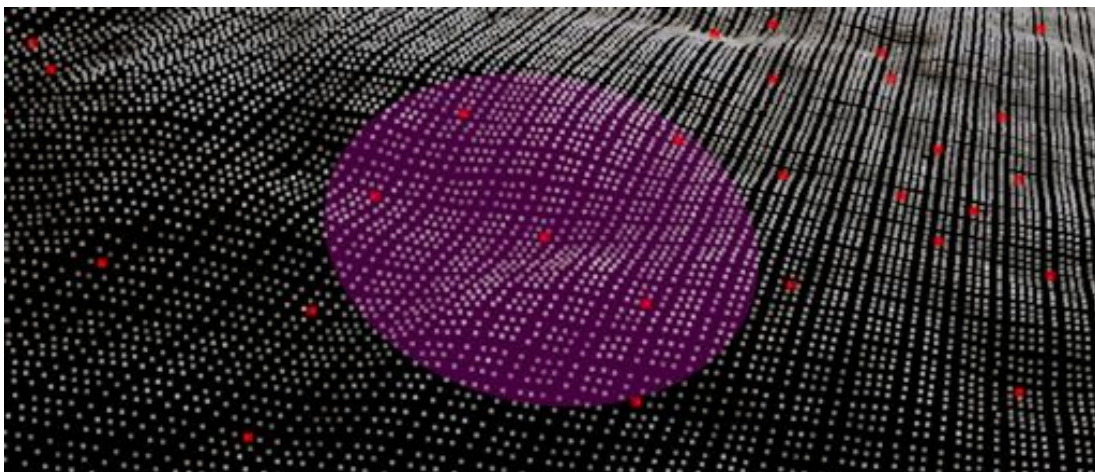


Figure 6.5: Spherical descriptor footprint with a radius of approx. 2mm. We observe that the peck mark is completely within the footprint.

at local maxima in the depth map and the depth has a value of zero at the image plane. We observe, that almost each peck mark has a point marked as peak at its lowest point. Hence, we assume that these points are good center points around which we can describe the surface properly.

FPFH describes the local neighborhood of a point based on the angular differences between the surface normals of the point and its neighbors (see Figure 5.2). We propose a new descriptor *TopoFeat* based on the Euclidean distance between surface points and the angular differences of their normals. PFH is proposing the usage of the pairwise Euclidean distances between the surfels around the query point, too. But this is not included in the reference implementation due to performance reasons.¹ However, the main difference between PFH and the *TopoFeat* descriptor is the fact, that we are not considering pairwise differences of *all* points in the neighborhood of the query point but only the pairwise differences *between the query point and all points within the descriptor footprint*. The *TopoFeat* descriptor is constructed as follows:

1. Select a query point (the point we want to construct the description for) and a radius r for the neighborhood that should be considered.
2. Select all points in the neighborhood.
3. Calculate the Euclidean distance and the angular difference of the normals between the query point and each point in the neighborhood. The resulting feature vector contains a quadruple for the difference between the query point and each point within the footprint (see Figure 6.5) consisting of three angular differences and the Euclidean distance between the two points (see Figure 5.2).
4. Sort the quadruples by the Euclidean distance.
5. Select the number of concentric shells m the spherical descriptor footprint should be divided in.
6. Divide r evenly in m parts.
7. Put each quadruple in one of the m parts according to its Euclidean distance to the query point and discard the distances. This results in m sets of triples containing the angular differences between the query point and all points in one of the m concentric shells.
8. Select a number of bins l for a 3D histogram.
9. For each of the m parts, aggregate the values of the triples in the 3D histogram. This results in a 3D histogram with l^3 bins for each of the m concentric shells. The m histograms contain information about the gradients of the surface in each of the m shells. The resulting feature vector has a size of $m \times l^3$.

We do not extract the *TopoFeat* densely for each point but only at points that are at the bottom of peck marks (see Figure 6.5). We assume, that the surface description is sensitive to the size of the peck mark(s) within the descriptor footprint for which it is computed. Hence, we assume that the descriptor allows us to distinguish different pecking styles.

¹See http://pointclouds.org/documentation/tutorials/pfh_estimation.php.

6. SIMILARITY OF PECKING STYLES

We extract the proposed TopoFeat at peak positions (i.e. the lowest points in peck marks, see Figure 6.5) and cluster the resulting features in 2,3,4 and 5 clusters and perform visual inspection to assess whether the descriptor allows to model peck style similarity well.

6.3 Results

In this section we present qualitative results of peck style clustering based on 3D description of the surface. Due to the computational complexity and the preliminary status of point cloud-based peck style analysis we employ nine 3D scans for our evaluation, namely scans 2, 3, 8, 9, 10, 19, 21, 24 and 26 (see Appendix A for a description of the dataset). We chose the set of scans with the aim to include figures that are pecked homogeneously and figures that are pecked in visually different pecking styles.

6.3.1 Dense sampling

We extract the FPFH descriptor from all nine scans with radii of $r=1$, 2 and 5 mm. Due to its computational complexity we extract PFH from six scans (3, 8, 10, 19, 21 and 24) which we select out of the nine. The selection criterion is the quality of the clustering of the FPFH descriptors. We choose the three point clouds for which we achieved good results and the three point clouds for which we achieved bad results. We use radii of $r=1$ and 2 mm as descriptor computation with $r=5$ mm did not terminate in a reasonable time. We show examples for good and for bad results.

Figures 6.2 and 6.3 show two of the point clouds we use for qualitative analysis containing distinguishable pecking styles. The warrior in Figure 6.2 contains two visually easily distinguishable pecking styles which are in clearly distinguishable regions. Furthermore, the head of the warrior shows a peck mark density between the two styles of the body and the superimposed arrow. We assume that this scan can serve as a basic test whether 3D descriptors can distinguish different pecking styles. The second scan we employ (see Figure 6.3) poses a harder challenge. The pecking styles in the figure at the upper left part of the scan are interwoven, possibly as result of superimposition of large pecks over a finely pecked warrior figure. Figure 6.6 shows a point cloud with a homogeneous pecking style. Figure 6.7 shows the results of pecking style clustering with 2, 3 and 4 clusters. We observe that the clustering with two clusters (upper right figure) separates the body and the arrow while being ambiguous for the head and the foot at the lower left of the figure. The ambiguity results in speckled regions. The results with three and four clusters do not improve the separation of the differently pecked regions.

Figure 6.8 shows the results of pecking style clustering with 2 and 3 clusters as well as a detail of the warrior figure in the scene. In the clustering with two clusters (upper right) we observe that the superimposition over the warrior is clearly distinguishable while the distribution of the two clusters in the other parts of the scan remains unclear. The clustering with three clusters (lower left) shows a comparable distinction for the warrior figure, the green cluster in the clustering with two clusters is split into a green



Figure 6.6: A petroglyph with homogeneous pecking style.

6. SIMILARITY OF PECKING STYLES



Figure 6.7: Visualization of clustering results of densely extracted FPFH descriptors with $r=5\text{mm}$. The figure in the upper left is the original point cloud, the other three figures are the results with 2, 3 and 4 clusters respectively. We observe that the clustering with two clusters is modeling the two visually observed pecking styles well. The clusterings with 3 and 4 clusters do not show improvements.

and a blue cluster. The details at the lower right expose that the superimposition over the warrior is modeled very well in both cases.

Figure 6.9 shows an example for a petroglyph with homogeneous pecking style. We observe that the clustering algorithm clusters the data in the required number of clusters, as it is a property of k -means to always create a clustering with the requested number of clusters. But, we observe that the clustering shows a pattern comparable to vertical stripes.

Figure 6.10 shows a comparison of the two descriptors (rows) we use and different footprints (columns) for descriptor extraction. We observe that the larger the footprint (it increases from left to right) the better the results we achieve. The two pecking styles are separated optimally with FPFH with $r=5\text{mm}$. We assume that the very large peck marks the superimposition is consisting of are described well with descriptor footprints with a diameter of 1cm as the best result is surprisingly achieved with FPFH which contains only a fraction of the surface topology information compared to PFH which we couldn't extract with the largest footprint due to computational complexity. Please refer to Subsection 5.1.2 for a detailed description of the difference between PFH and FPFH.

Figure 6.11 shows a surface detail. We observe that the flatter surface areas with shallow peck marks and the rougher surface areas with large peck marks are separated well.

6.3.2 Topology-based sampling

We extract the topology feature with a radius $r=2\text{mm}$ from all nine scans. We show examples for a good and for a bad result.

Figure 6.12 shows the result of the computation and clustering of our TopoFeat. We observe that the distinction between body and arrow is modeled well. This is especially remarkable, as the number of peak positions used for feature extraction is around one percent compared with the number of points used for dense extraction. Thus the computational demand for describing and clustering the pecking styles is roughly only ne percent compared to the dense approach. The main restriction we faced during our experiments was the computational demand of 3D descriptors. Hence, our results with TopoFeat demonstrate that the utilization of peak positions as salient points clearly improves the feasibility of using a 3D point cloud for similarity estimation of pecking styles.

Figure 6.13 shows the result of another example using our proposed TopoFeat. We observe that the superimposition is only distinguishable partly. This scan which has been separated well with the dense approach (see Figure 6.8) shows the limitations of the current status of our TopoFeat.

6. SIMILARITY OF PECKING STYLES

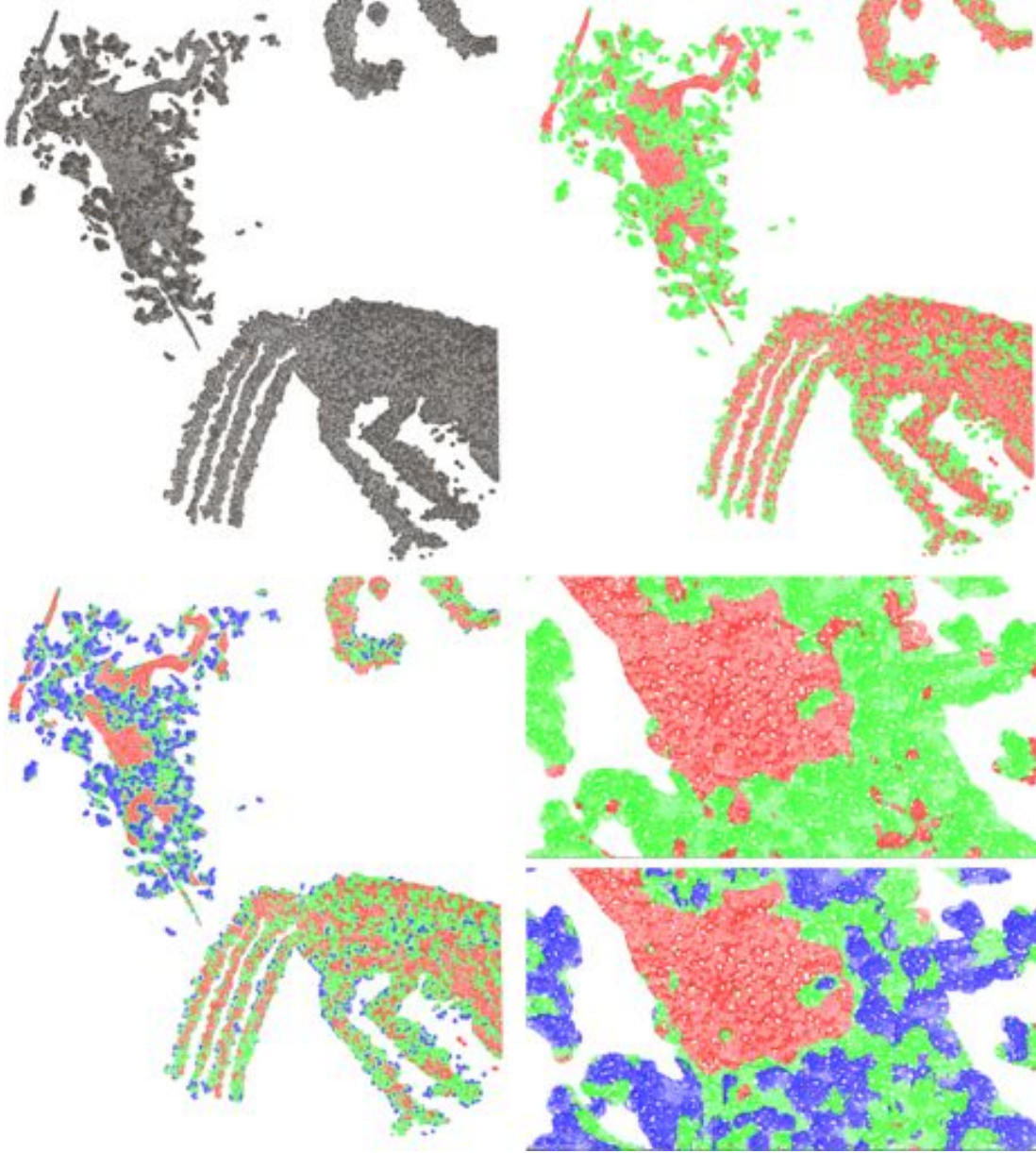


Figure 6.8: Visualization of clustering results of densely extracted FPFH descriptors with $r=2\text{mm}$. The figure in the upper left is the original point cloud, the upper right figure shows the result of clustering with two clusters while the figure at the lower left shows the result with three clusters. The two details at the lower right of the figure show the central part of the warrior in the scene. We observe that the superimposition on the figure is clearly distinguishable in the two and three cluster case.

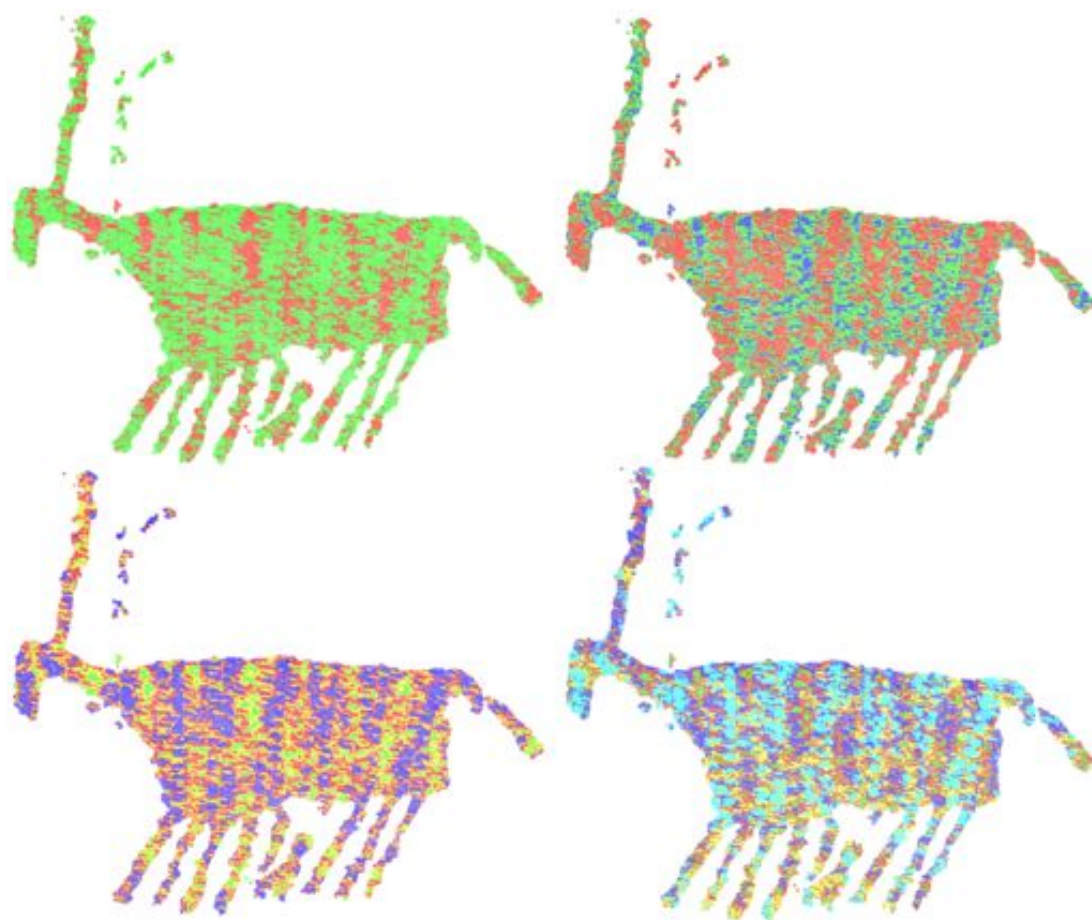


Figure 6.9: Visualization of clustering results of densely extracted FPFH descriptors with $r=2\text{mm}$. We show results with 2,3,4 and 5 clusters. We observe that the clustering results in vertical stripes that are assumably not caused by the surface topology of the homogeneous pecking style of the petroglyph (see Figure 6.6).

6. SIMILARITY OF PECKING STYLES



Figure 6.10: Comparison of different descriptors and footprints. The first row shows clustering results based on the FPFH descriptor with $r=1, 2$ and 5mm (from left to right). The second row contains clustering with PFH with $r=1$ and 2mm (from left to right). We observe, that this scan requires a large descriptor footprint as the best result is FPFH with $r=5\text{mm}$. The original scan is shown in Figure 6.2.

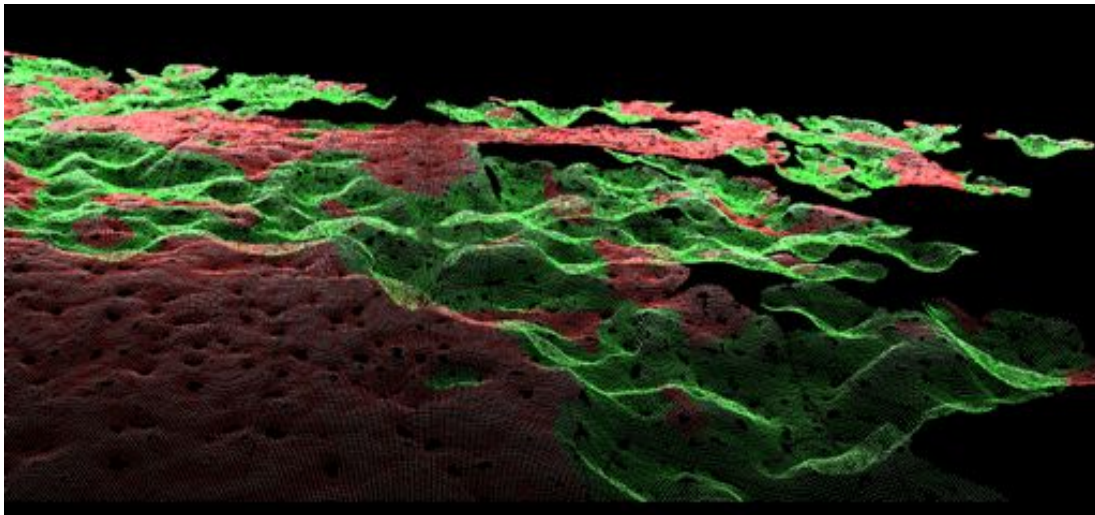


Figure 6.11: A surface detail of a point cloud the points of which have been clustered into two clusters based on PFH with $r=2\text{mm}$. We observe that the flat regions in the surface are in one cluster while the deep peck marks are in the other cluster.

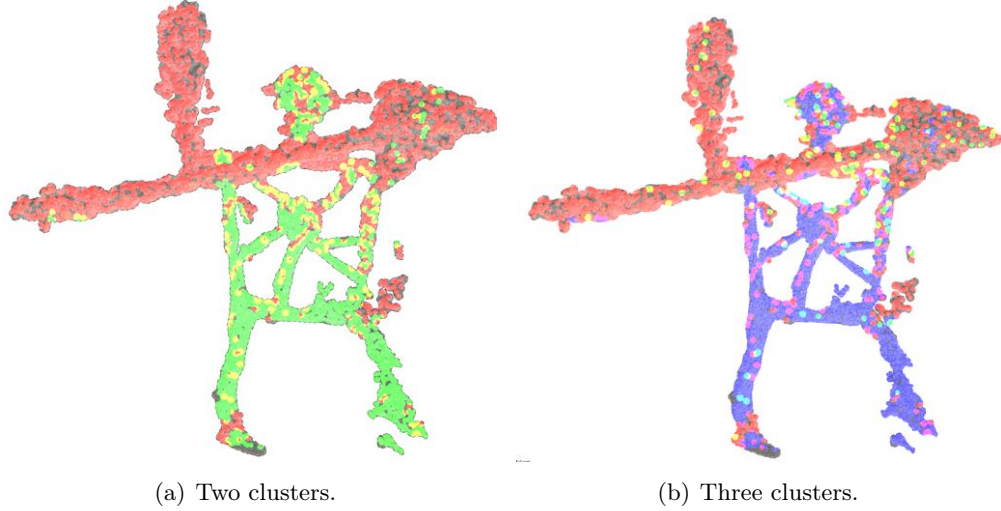


Figure 6.12: Warrior with surface spheres around peaks serving as basis for topological feature extraction. We observe, that the different pecking styles are modeled well (See Figure 6.2). The third cluster (green) in b) is very small. Please note, that the circular region colored around a peak has only a quarter of the radius that is used as query point neighborhood for feature calculation. We reduced the radius for visual clarity. Overlapping areas of differently colored circles show a mixture of the colors, e.g. yellow in the case of red and green in Fig b).

6.4 Conclusion

We have investigated the similarity of pecking styles using different types of sampling (dense and topology-based). We have shown that 3D surface descriptors can model pecking styles well in several cases. In other cases, the pecking style clustering fails.

The main restriction is the computational demand which prevents extensive experimentation. Hence, future work should focus on the description of the surface around salient points as this approach reduces the computational demand drastically. Additionally, in our experiments we experienced that the evaluation of pecking style clustering is particularly challenging due to the absence of ground truth information. An unsupervised clustering can only be as good as the underlying descriptors. Without additional knowledge about the searched patterns (pecking styles) it is not possible to select adequate parameters for the clustering and to remove irrelevant or even misleading parameters which can corrupt the clustering easily. In an unsupervised setting the selection of adequate parameters remains a task that has to be performed manually by the user and hence prohibits systematic evaluation.

We recommend the use of a semi-supervised approach in future, where the user provides hints to a machine learning system. In this way expert knowledge can be integrated and a suitable set of parameters for a given pecking style can be selected in a supervised fashion. This could facilitate the generation of computational models for

6. SIMILARITY OF PECKING STYLES

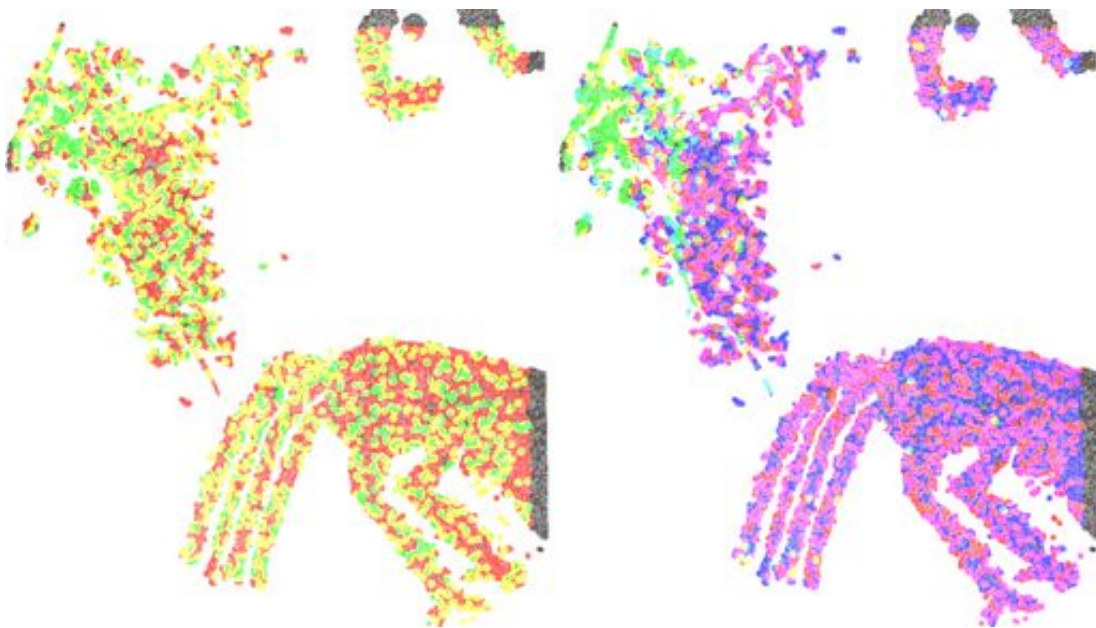


Figure 6.13: A complex scene with surface spheres around peaks serving as basis for topological feature extraction. We show results with 2 and 3 clusters. We observe, that the different pecking styles are modeled only partly (See Figure 6.3). Please note, that the circular region colored around a peak has only a quarter of the radius that is used as query point neighborhood for feature calculation. We reduced the radius for visual clarity. Overlapping areas of differently colored circles show a mixture of the colors, e.g. yellow in the case of red and green in the left figure.

different pecking styles. The usage of automatic feature selection during model learning (e.g. by sparse SVM or Random Forests) would enable to take the complete set of investigated parameters into account which in turn maximizes the number of pecking style attributes that can be considered in the analysis.

6. SIMILARITY OF PECKING STYLES

Description of Petroglyph Shapes

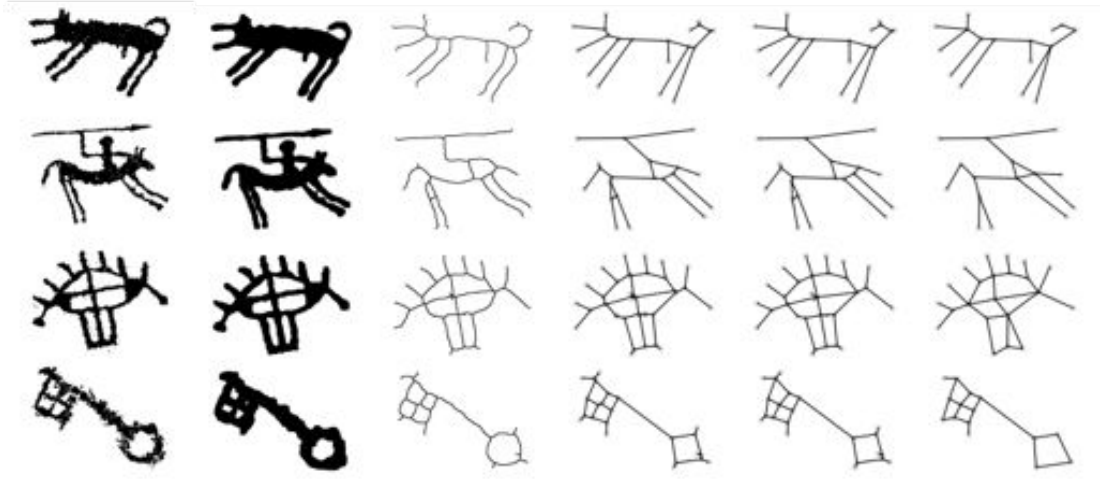


Figure 7.1: Petroglyphs with the corresponding smoothed figures, skeletons and the derived graphs.

IN this chapter we investigate the similarity of petroglyphs with respect to their shape. In contrast to the automated segmentation of rock surfaces containing petroglyphs, the automated classification of petroglyph shapes has attracted some research in the past years [ZWKL09][ZWKL11][DPdL12][DP13]. The large number of single petroglyphs makes the usage of automated analysis methods attractive. We propose and evaluate a graph-matching approach for shape similarity of petroglyphs (see Figure 7.1). The motivation to investigate the skeletal graphs of petroglyphs for shape matching is twofold. First, the material requires robustness against affine transformations and articulations. Second, we are not aware of existing petroglyph shape matching approaches using the skeletal graph.

7.1 Related Work

Numerous surveys about shape analysis have been published, comprehensive and important in this field are the surveys by Pavlidis [Pav78], Loncaric [Lon98], Zhang and

7. DESCRIPTION OF PETROGLYPH SHAPES

Lu [ZL04] as well as by Yang et al. [YKR08]. Classifications and taxonomies of shape descriptors have been proposed in the mentioned surveys in different variants. A widely used common denominator is the distinction between contour-based and region-based descriptors.

Latecki et al. compared several shape descriptors on the MPEG-7 CE-Shape-1 database [LLE00]. The database contains only complete shapes with closed contours. They propose three main categories for shape descriptors: Contour-based descriptors, region-based descriptors and skeleton-based descriptors. They investigated robustness to scaling and rotation, similarity-based retrieval as well as motion and non-rigid deformations. The weakest performing descriptor in all cases was the skeleton-based approach, the most significant drawback is the lack of robustness against scaling and rotation. The authors assume, that none of the existing approaches to compute skeletons is robust enough. But, since the publishing of this paper in the year 2000, promising skeletonisation algorithms have been proposed and evaluated (e.g. [BLL07]). We summarize region-based, contour-based and skeleton-based approaches that are relevant for our work and include petroglyph-related work where available.

7.1.1 Region-based Descriptors

Zhu et al. propose the usage of a slight modification of the generalised Hough transform (GHT) for the mining of large petroglyph datasets [ZWKL11]. The main arguments for GHT and against other shape similarity measures are the existence of petroglyph images where a single petroglyph consists of several parts and the possibility of merged parts of petroglyphs that drastically change the topology of the petroglyphs. They extensively evaluate their approach and achieve good results. However, they mostly evaluate synthetic petroglyph shapes or simple petroglyph shapes drawn by humans rather than the more exact tracings based on peck marks which we use (see Subsection 7.2.1). Deufemia et al. use the radon transform of petroglyph images as shape descriptor for unsupervised recognition via self-organizing maps (SOM) [DPdL12]. In a second step, they use a fuzzy visual language parser to solve ambiguous interpretations by incorporating archaeological knowledge. They evaluate the approach on a large dataset and achieve good results. However, Deufemia et al. as well as Zhu et al. do not consider partial or merged petroglyphs, i.e. part-based retrieval that is necessary for petroglyphs in real-world scenes.

Krish and Snyder propose the shape recognition approach SKS [KS08] which is based on the generalised Hough transform. They compare the performance of SKS with Hu moments, curvature scale space (CSS, see Subsection 7.1.2) matching and shape context (SC, see Subsection 7.1.2). Besides affine transformations, they evaluate partial shapes. The SKS feature performs good on partial shapes. But, the evaluation data set consists of 31 different shapes only and does not contain merged shapes.¹ Generally, region-based descriptors have the advantage that they do not need complete contours for descriptor extraction.

¹The data set does not include petroglyphs.

7.1.2 Contour-based Descriptors

Mokhtarian et al. propose curvature scale space (CSS) image matching [MM86][Mok95][MAK⁺97]. They smooth a contour by convolution with a Gaussian kernel in different scales (i.e. different kernel sizes of the Gaussian kernel). Subsequently, they find the curvature zero crossings on the contour. The descriptor - the CSS image - consists of the zero crossings in a diagram where the size of the Gaussian kernel is on the y-axis and the normalized path length of the curve is on the x-axis. This CSS image is used to match shapes. Mai et al. use the CSS descriptors to acquire contour segments invariant to affine transformations [MCH10]. They utilize the local maxima in the CSS image to locate the affine-invariant points and segment the contour at these points. They match the segments with a dynamic programming approach. They achieve very good experimental results. They outperform the dynamic programming approach by Petrakis et al. [PDM02], who are utilizing contour segments as well. Belongie et al. propose the widely used shape context (SC) descriptor [BMP02]. They sample points on the contour of an object. For each point, they compute the shape context based on the spatial distribution of the other points on the shape contour. They match two shapes by estimating the best transformation from one shape to the other and determining a dissimilarity based on shape context distance, appearance cost and transformation cost.

Deufemia et al. [DP13] propose a two stage classification of petroglyphs. They use shape context descriptors to provide an initial raw clustering with self-organizing maps. In the second step, they use an image deformation model to classify the petroglyph shapes. They evaluate their approach on a relatively small dataset (17 classes with 3 exemplars) that they enlarge by using 30 affine transformations of each image.

7.1.3 Skeleton-based Descriptors

Siddiqi and Kimia propose a shock grammar for recognition [SK96]. Later, Siddiqi et al. propose the usage of shock graphs for shape matching [SSDZ99]. Shocks are used to provide a structural description of 2D shapes. They are contour-based and deliver a medial axis of the shape, that has additional information for each part of the skeleton. The representation of the shape is a directed acyclic shock graph, which is used for shape matching.

Aslan Skeletons are coarse skeletons [AEET08]. They are matched via tree edit distance and have been evaluated on different data sets with good results [BET09][ET10]. However, while they are insensitive to articulation and affine transformations, they can not be used for merged shapes, for shapes with holes and for shapes with large missing parts.

Ling and Jacobs propose the usage of the inner-distance, which is the shortest path between two landmark points (in this case contour points are used) of a shape [LJ07]. Hence, they implicitly embed skeletal information in the descriptor. The distance between two contour points is the shortest path within the silhouette of the shape instead of the Euclidean distance between the two points. They use the idea for three approaches to shape description. First, they combine the inner-distance with multi-dimensional scal-

7. DESCRIPTION OF PETROGLYPH SHAPES



Figure 7.2: This figure shows a part of a tracing of a rock in Valcamonica (Coren di Redondo, Rock 1). ©Alberto Marretta, used with permission.

ing. Second, they utilize the inner-distance to build a new descriptor based on shape context, and third they extend the second approach with appearance information of the shapes along the inner-distance lines. They evaluate the approach on several datasets with good results. They state, that the proposed descriptors are invariant/insensitive to articulation and are capable to capture part structures.

Bai and Latecki introduce a skeleton-based approach that matches silhouettes based on skeleton paths, which are the geodesic paths between skeleton endpoints [BL08]. The shortest paths are represented by the radii of the maximum inscribed discs at skeleton points. They use DCE (Discrete Curve Evolution [LL00]) for the skeleton pruning. The descriptor is on two layers. First, the description of a skeleton endpoint is constructed from the shortest paths starting at this point, and second the similarity of two shapes is computed by matching the descriptors of the skeleton endpoints. They

experimentally show, that the method is robust against articulations, stretching and contour deformations. Bai et al. combine contour features with skeletal features to improve shape classification [BLT09]. They state, that contour-based approaches can represent detailed information well, and are up to a certain extent robust against partial and merged shapes but lack invariance against articulation and non-rigid deformation. In contrast to that, skeleton-based approaches are robust against non-rigid deformations. For the contour segments, they follow the ideas of Sun and Super [SS05], but use DCE to determine the segments. They achieve 96.6% classification rate on the MPEG-7 CE-Shape-1 [LLE00] database. Xu et al. extend the skeleton path approach [BL08] by also considering junction points for the skeleton paths descriptor [XWLB10]. They call the junction points and end points of the skeleton graph critical points. They merge junction nodes based on the paths from the junction nodes to the end nodes. If the sum of path distances of two nodes normalised by the number of end nodes of the graph is below a set threshold the two nodes are merged. They achieve slightly better results than Bai and Latecki [BL08], and state, that the method is efficient even in the presence of articulation as well as partial and merged shapes. The line of work summarized in this paragraph is mostly built on the skeleton pruning algorithm based on DCE proposed by Bai et al. [BLL07].

7.1.4 Discussion of Shape Descriptors

All existing approaches dealing with petroglyphs have drawbacks. [ZWKL11] is evaluated at large scale on publicly available data, but the data are different from the material we are using. They use *inter alia* sketch data and synthetic petroglyph shapes. In contrast to that, our data is the product of the standard archaeological documentation process which is manual tracing of individual peck-marks on transparent sheets (see Section 2.6). [DPdL12] and [DP13] evaluate on small datasets. [DP14] have a large dataset and compare three different descriptors in their evaluation, but they don't evaluate the combination of descriptors. Moreover, they only compare region-based descriptors. In our approach, we aim to overcome these shortcomings.

7.1.5 Graph Matching

To our knowledge, there is no work on petroglyphs that utilizes skeletons or skeletal graphs. We aim at investigating the skeletal graph for petroglyph similarity modeling. In our approach, we use the promising algorithm by Bai et al. [BLL07] for skeletonisation and skeleton pruning. To investigate the distinctiveness of skeletal graphs, we model petroglyph similarity as graph similarity. In the following, we summarize works that use graph matching to model shape similarity and popular graph matching approaches. There are numerous approaches in shape matching that utilize graphs. Comparable to our approach are methods, that model shape similarity as similarity of the skeletal graph. Klein et al. use tree edit distance to match shapes described by their shock graphs [KSK01]. Di Ruberto uses attributed skeletal graphs to model shape similarity [DR04]. Aslan skeletal graphs are matched with tree edit distance

7. DESCRIPTION OF PETROGLYPH SHAPES

[BET09][ET10]. For our material (see Section 7.2.1), tree matching is not sufficient, as the skeletal graphs may contain cycles. Hence, we have to use graph matching. There is a long line of work in graph matching. A recent volume of workshop proceedings edited by Kropatsch et al. includes papers on many aspects of graph-based representations in pattern recognition [KAH⁺13].

A popular and intuitive similarity measure for a pair of graphs A and B is the graph edit distance (GED). GED defines similarity of two graphs as the minimum number of edit operations (remove node, add node, remove edge, add edge) that are needed to transform graph A to graph B . The computation of the graph edit distance is NP-hard [ZTW⁺09]. This problem is addressed in several ways. There are approximations for the GED, e.g. the widely used Hungarian algorithm, or A* beam search. Another way is to use graph spectra (e.g. [DvLV08]) or to embed node and edge attributes of graphs in vector spaces and subsequently use standard similarity measures and machine learning methods to match similarity (e.g. [GVB12] or [LSYZ11]).

7.2 Proposed Approach

7.2.1 Benchmark Dataset

In this chapter we use a fraction of our dataset described in Subsection 2.6.2 consisting of one hundred tracings of individual petroglyphs in ten classes. Figure 7.3 shows examples of our material. Note the high intra-class variability not only in terms of affine transformations and articulations but also in general differences of the shapes in a class. Some classes are perceptually very close, e.g. classes two and three.

7.2.2 Overview of our Approach

The literature review suggests, that a combination of contour-based and skeletal descriptors should yield optimum results. The existing petroglyph shape retrieval systems utilize region-based features [ZWKL11][DPdL12] as well as contour-based features [DP13]. These methods are the baseline reference methods we compare our approach with. We concentrate on the investigation of skeletal features for the petroglyph shape recognition problem, because the petroglyph shapes are intuitively already skeletal shapes. We propose to model similarity of petroglyphs as a skeletal graph matching problem. We derive the graphs from the skeletons of the petroglyph shapes. We use each end point and each junction point of a given petroglyph skeleton as nodes for our graph, and the skeleton branches connecting these points as edges. See Figure 7.4 for exemplary skeletons and derived graphs. We utilize the skeletal graph as descriptor that is invariant to affine transformations as well as articulations. Hence, we discard all spatial information, as we are only interested in topological information. We match the resulting undirected graphs following two strategies. First, we utilize the graph edit distance (GED) as pairwise similarity measure, and second, we use graph embedding (GE) to create a feature vector for each graph and match these feature vectors. We embed the



Figure 7.3: This figure shows the petroglyph dataset that we investigate in this chapter. Each column contains examples of one class. We observe, that some of the classes have high intra-class variance of shape. The petroglyphs are from various rocks in Valcamonica. ©CCSP - Centro Camuno di Studi Preistorici and Alberto Marretta, used with permission.

7. DESCRIPTION OF PETROGLYPH SHAPES

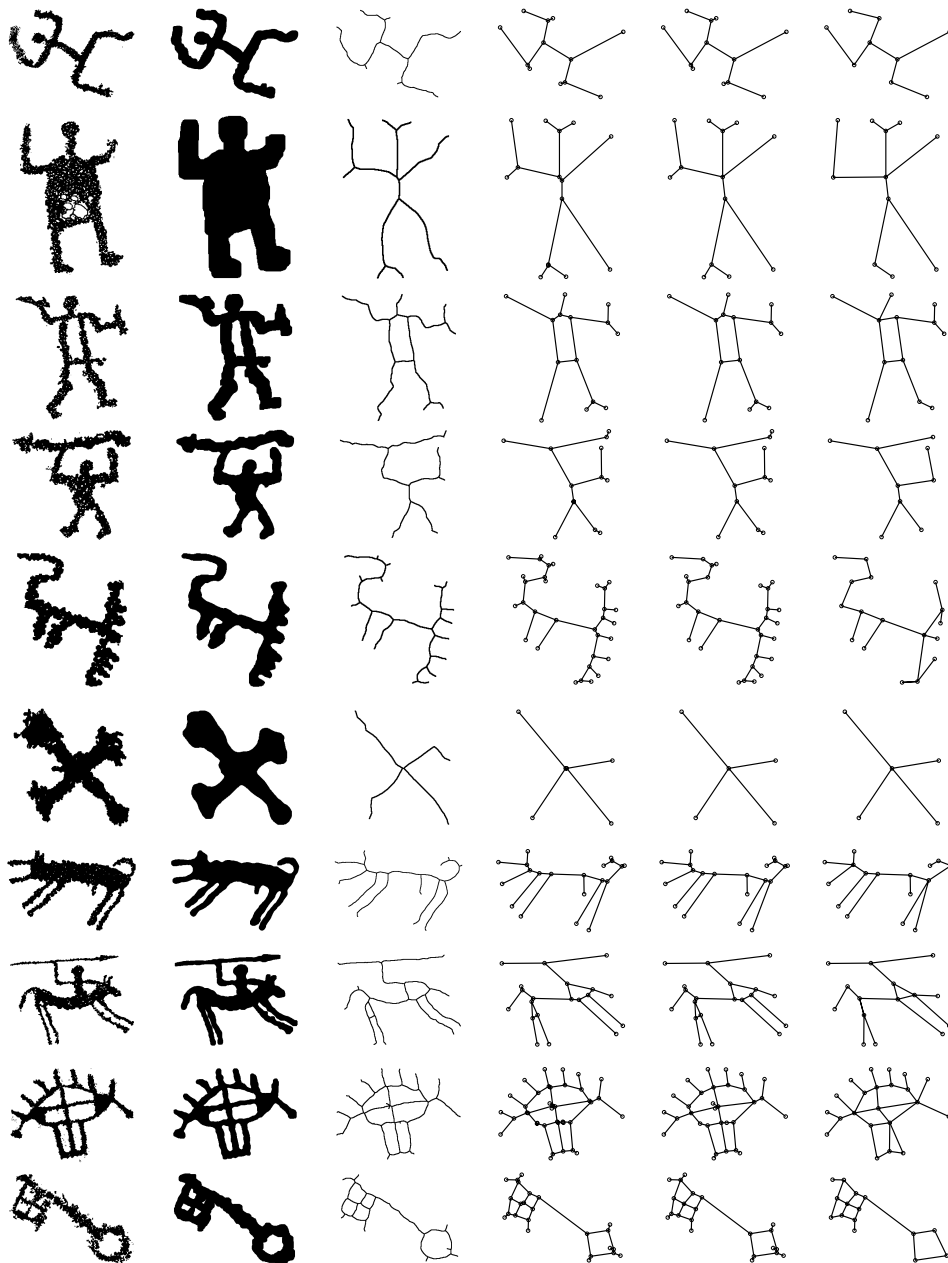


Figure 7.4: This figure shows petroglyph tracings (column 1), pre-processed shapes (column 2), extracted skeletons (column 3) and the derived graphs with a pruning threshold of 2, 10 and 30px (columns 4-6). We observe, that the employed skeletonisation algorithm fails to extract details in some cases (heads in rows 1 and 4), while in other cases small details are covered well. Furthermore, we observe, that a high pruning threshold removes skeletal noise (rows 1, 2, 4, 6, 9 and 10), but may also discard relevant topological information (row 5).

graphs in feature vectors of a normed length by extracting several graph properties (see Table 7.1), and calculate pairwise distances with Euclidean distance. We classify using a k-NN classifier with an extension for intermediate descriptor fusion.

7.2.3 Descriptors

We compare the distinctiveness of contour-based as well as region-based shape descriptors with the distinctiveness of undirected skeletal graphs for the petroglyph classification problem. As baseline methods, we utilize a) Shape Context (SC) [BMP02], which has been used by Deufemia et al. for petroglyph classification [DP13], b) Inner Distance Shape Context (IDSC) [LJ07], as well as c) General Hough Transform (GHT) proposed by Zhu et al. for petroglyph classification [ZWKL11].¹ Our proposed petroglyph descriptor makes use of the skeletal graphs in two ways. First, we use the graphs directly as descriptors and measure the similarity with GED. We employ the A* algorithm with beam search and the Hungarian algorithm. Both variants tolerate cycles in the graphs.² Second, we utilize topology features of the undirected graphs for GE.³ Please refer to Table 7.1 for the list of selected features for petroglyph description. We evaluate the distinctiveness of single topology features and of combinations of topology features. In order to be able to use skeletonisation and contour-based descriptors, we need to pre-process our material to achieve continuous boundaries. We use standard filters and morphological operations for this purpose. The resulting shape images are the input for SC, IDSC as well as for the skeletonisation for which we use the method by Bai et al. [BLL07].⁴ We normalize the width of the input petroglyph to 500px and prune the skeletons by joining nodes which have a spatial distance smaller than a threshold. The threshold varies between 2 and 30px. See Figure 7.4 for accordingly pruned graphs.. GHT is computed on the original petroglyph images.⁵

7.2.4 Descriptor Fusion and Classification

Additionally to the performance of individual descriptors, we evaluate the performance of descriptor combinations. For SC, IDSC, GHT and GED we compute pairwise distances. For the feature vectors we obtain by GE, we calculate pairwise Euclidean distances. The combination of descriptors via the combination (i.e. unweighted summation) of distance matrices would require a normalization in all utilized feature spaces, which could only be provided by heuristically determined thresholds for the maximum dissimilarity that can occur for one specific descriptor. We want to avoid this step, and aim at preserving complementary distinctiveness as far as possible in the classification

¹For all three descriptors, implementations have been kindly provided by the respective authors.

²We use the implementation in the Graph Matching Framework kindly provided by Kaspar Riesen [REB13].

³We utilize the MIT strategic engineering tools for network analysis kindly provided by Bounova and de Weck [BdW12]: http://strategic.mit.edu/downloads.php?page=matlab_networks.

⁴We thank Bai et al. for providing the implementation.

⁵We use resolutions of 10x10px, 20x20px and 30x30px. We report on the best performing resolution 10x10px.

7. DESCRIPTION OF PETROGLYPH SHAPES

Table 7.1: Topology based features extracted from the skeletal graphs. Each feature is a description of the whole graph. Please refer to [BdW12] for more details.

ID	Feature	Description
1	numNodes	Number of nodes
2	numEdges	Number of edges
3	numCycles	Number of independent cycles, also known as the cyclomatic number of a graph
4	linkDensity	Link density, i.e. ratio of existing links to maximum possible links
5	avgDegree	Average number of links over all nodes
6	numLeaves	Number of leaves, i.e. number of nodes with only one link
7 - 11	histDegrees	5 bin histogram of node degrees, i.e. counts of all nodes with 1,2...5 links. The maximum occurring degree in our dataset is 5
12	sMetric	Sum of degree products across all edges, i.e. for each edge, multiply the degrees of the two nodes connected by the edge and finally sum up the products
13	graphEnergy	Sum of the absolute values of real components of the eigenvalues
14	avgNeighDegree	Average of the average neighboring degrees of all nodes
15	avgCloseness	Average of closeness over all nodes
16	pearson	Pearson coefficient for the degree sequence of all edges of the graph
17	richClub	Rich club metric for all nodes with a degree larger than 1
18	algebConnect	Algebraic connectivity, i.e. the second smallest eigenvalue of the Laplacian
19	diameter	The longest shortest path between any two nodes in the graph
20	avgPathlength	Average shortest path
21	graphRadius	Minimum vertex eccentricity

Table 7.2: Classification accuracy of GED in percent. We use LOOCV-validated 5-NN classification. The maximum number of open paths for the A* beam approximation is 6000.

	Pruning threshold					
	2	10	15	20	25	30
Hungarian	52	57	57	30	39	37
A* beam	51	57	49	47	47	47

process. Instead of merging the similarity matrices, we employ single k-NN classifiers ($k=5$) for each individual feature. A straightforward approach would be to combine the classification results of all classifiers by majority voting. To obtain a richer and more expressive basis for making a decision, we refrain from this simple combination of classification results. Instead, we fuse the classifiers at an intermediate step. We use the class labels of the five nearest neighbors of each descriptor and concatenate these to a set of nearest neighbors which contains $5n$ class labels, with n being the number of descriptors combined. Subsequently, we classify according to the conventional rule of k-NN with a majority vote. This incorporates more information than a voting based on the classification results, as we implicitly include the probability of the classification result in form of the number of class members among the nearest neighbors. We validate our results of single descriptors as well as fused descriptors by leave-one-out-cross-validation (LOOCV). We employ accuracy as quality measure, i.e. the ratio of the number of correctly classified petroglyphs to the total number of classified petroglyphs.

7.3 Experimental Results

7.3.1 Graph Matching

Table 7.2 shows the results of GED employed on the unweighted undirected skeletal graphs. The maximum accuracy is 57%. Both employed methods achieve this result at a pruning threshold of 10px. We observe, that the results of both methods tend to decrease with an increasing pruning threshold larger than 10px. We assume, that the distinctiveness of the skeletal graphs first increases, as skeletal noise is pruned, and then decreases after a maximum of 10px as more and more topological distinctiveness is removed in the pruning process (see Figure 7.4). Furthermore, we observe, that for higher pruning thresholds, A* beam search outperforms the Hungarian algorithm.

Table 7.3 summarizes the results of the evaluation of single topological features of GE. We observe, that feature 13 (graph energy) yields the best result with 47% accuracy given a pruning threshold of 10px. We achieve also the second and the third best results with a pruning threshold of 10px. This confirms the GED results (see Table 7.2), where a pruning threshold of 10px delivers the best results as well. Furthermore, we observe,

7. DESCRIPTION OF PETROGLYPH SHAPES

Table 7.3: Classification accuracy (in percent) of GE utilizing single scalar topological features. We use LOOCV-validated 5-NN classification. p denotes the size of the pruning threshold. The three best values are emphasized. Please refer to Table 7.1 for the descriptions corresponding to the feature IDs.

	Feature ID																					
p	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Avg.
2	25	21	14	35	27	40	40	10	33	10	10	22	25	21	30	23	35	30	27	29	14	25
10	37	44	14	29	31	40	40	9	25	7	10	33	47	23	43	23	29	32	25	31	20	28
15	35	25	17	38	33	37	37	5	21	7	10	30	37	16	30	22	38	30	23	39	23	26
20	31	28	18	31	37	21	21	8	16	10	10	35	35	15	33	26	31	33	29	37	21	25
25	33	32	20	37	39	25	25	12	14	10	10	25	39	13	37	17	37	37	26	33	29	26
30	27	36	13	28	35	23	23	13	8	10	10	33	27	9	36	22	28	34	29	26	28	24
Avg.	31	31	16	33	34	31	31	10	20	9	10	30	35	16	35	22	33	33	27	33	23	

that features 8,10 and 11 (bins 2,4 and 5 of the degree histogram) perform around random (10%). We assume that this is due to the fact, that most graph nodes that are not leaves seem to be 3-connected (see Figure 7.4). Hence, the counts of 2, 4 and 5-connected nodes have weak discriminative capabilities.

Table 7.4 reports the GE performance of the best feature combination of each dimension from 1 to 10. We employ brute-force feature selection for each number of feature dimensions, i.e. we evaluate all possible combinations for 1 to 10 out of the 21 single features (see Table 7.1). We observe, that feature combination strongly improves results. The maximum accuracy is 57%, as it is the case with GED as well.

7.3.2 Comparison with Baseline Descriptors and Descriptor Fusion

Table 7.5 contains the results of shape descriptors with which we compare our skeletal graph approach as well as combinations thereof. We observe, that SC and IDSC perform better than GE when employed as single descriptors. GE clearly outperforms the dedicated petroglyph descriptor GHT of [ZWKL11]. Descriptor fusion generally improves results. The fusion of two or three shape descriptors improves the results slightly. The fusion of all four descriptors improves the results from 82% for the best single descriptor to 88%. This demonstrates, that the descriptors contain complementary information that is well preserved by using the proposed k-NN fusion method.

To discuss the limitations of our approach and the weak performance of GHT on our dataset compared to the datasets employed by Zhu et al. [ZWKL11], we present example query petroglyphs and their nearest neighbors. Figure 7.5 shows a query petroglyph, that is misclassified by GED as well as by GHT. We observe, that the skeletal graph of the query antropomorph figure has a topology which is similar to the topology of a cross. Hence, the five nearest neighbors are crosses. The nearest neighbors computed by GHT also fail to determine the correct class for the query petroglyph. We observe, that

Table 7.4: Classification accuracy (in percent) of GE using combinations of 1 to 10 single topology features that perform best for each dimension. We use LOOCV-validated 5-NN classification with brute-force feature selection. The pruning threshold is 10px. The best values are emphasized. Please refer to Table 7.1 for the descriptions corresponding to the feature IDs.

Feature IDs	Accuracy
13	47
2, 15	54
4, 6, 18	56
6, 7, 18, 20	57
6, 7, 15, 18, 20	57
4, 6, 7, 15, 18, 20	56
1, 2, 4, 6, 7, 15, 17	55
1, 2, 4, 6, 7, 13, 15, 17	55
1, 2, 4, 6, 7, 14, 15, 17, 20	53
1, 2, 4, 6, 7, 13, 14, 15, 17, 20	53

Table 7.5: Classification accuracy (in percent) of GE, SC, IDSC and GHT and combinations thereof. We use LOOCV-validated 5-NN classification for single shape features and the classifier fusion method described in Section 7.2.4.

Descriptor	Single								Fused							
GE	x				x	x	x	x	x	x						x
IDSC		x			x			x	x		x	x	x			x
SC			x			x		x		x	x		x	x	x	x
GHT				x			x		x	x		x	x	x	x	x
Accuracy	57	81	82	39	80	78	54	86	84	83	84	81	86	81	88	

7. DESCRIPTION OF PETROGLYPH SHAPES

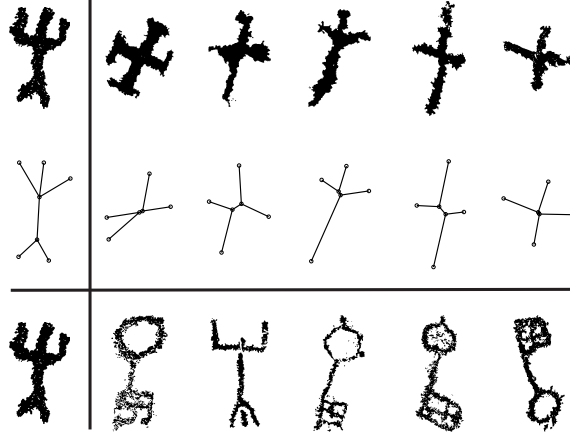


Figure 7.5: This figure shows a sample which is misclassified with GED and GHT. The query image is on the left, and the five nearest neighbors on the right. The first two rows show the result for GED and the utilized graphs. The third row shows the result for GHT.

the spatial distribution of the pixels in the query image is comparable to the nearest neighbors. Figure 7.6 shows a query petroglyph, that is correctly classified by GE and misclassified by GHT. We observe, that in the case of GE, the four nearest neighbors are topologically very close to the query graph. The fifth neighbor is different. But, we have to take into account, that GE matches with a set of features, that cannot necessarily be understood intuitively (see Table 7.4). The GHT nearest neighbors show comparable pixel distributions. We assume, that the less competitive performance of GHT on our dataset is related to the fact, that the test datasets used by Zhu et al. are manual transcriptions of petroglyph skeletons (or sometimes outlines, see [ZWKL11] p95) which leads to simpler shapes than the detailed tracings of peck marks in our material.

7.4 Conclusion

We present a novel petroglyph descriptor based on the skeletal graph topology and propose matching with graph edit distance (GED) and graph embedding (GE). For GE, we propose 21 different scalar topological features. We evaluate the descriptor and the matching on a petroglyph dataset containing 10 classes with 10 exemplars and compare the performance with other shape descriptors used in petroglyph classification.

Matching of the skeletal graphs with GE and with GED delivers comparable results. Both matching methods achieve 57% accuracy. GED is of high computational complexity, whereas GE has low computational demand due to low feature vector dimensionality. The two best performing combinations of topology features have only 4 and 5 feature dimensions. The contour-based features SC and IDSC outperform the region-based GHT and the skeletal graph-based GE and GED. GE and GED outperform the region-based GHT. The proposed descriptor fusion clearly improves results. In 5 of

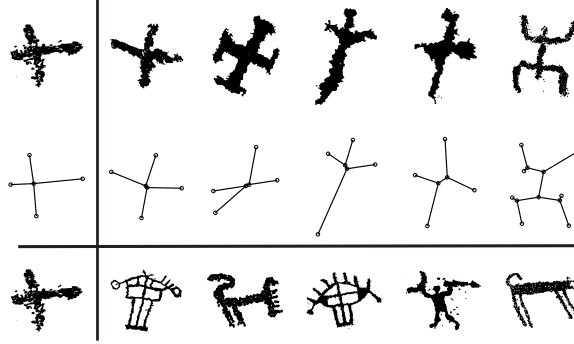


Figure 7.6: This figure shows a sample which is correctly classified with GE and misclassified with GHT. The query image is on the left, and the five nearest neighbors on the right. The first two rows show the result for GE and the utilized graphs. The third row shows the result for GHT.

7 descriptor combinations the usage of our descriptor improves results (see Table 7.5). The combination of our graph-based petroglyph descriptor with other descriptors yields a classification performance of 88% which is not achieved without the proposed skeletal descriptor. This shows that the skeletal features represent information not captured by the contour-based features and the region-based features. We conclude that descriptors derived from skeletons are valuable for petroglyph classification.

Future work could include improvement of our petroglyph descriptor based on the skeletal graph in two ways. First, we aim at improving pre-processing of the shapes as well as skeletonisation. Second, we could investigate the suitability of spatial features. In this chapter, we investigated topological information of the skeletal graph, which is a highly invariant abstraction of the skeleton. In future, we could investigate, whether spatial relations of graph parts contain valuable information for our task, because our material has already undergone one step of abstraction by the artists, who created the petroglyphs.

7. DESCRIPTION OF PETROGLYPH SHAPES

Skeletonisation of Petroglyph Shapes

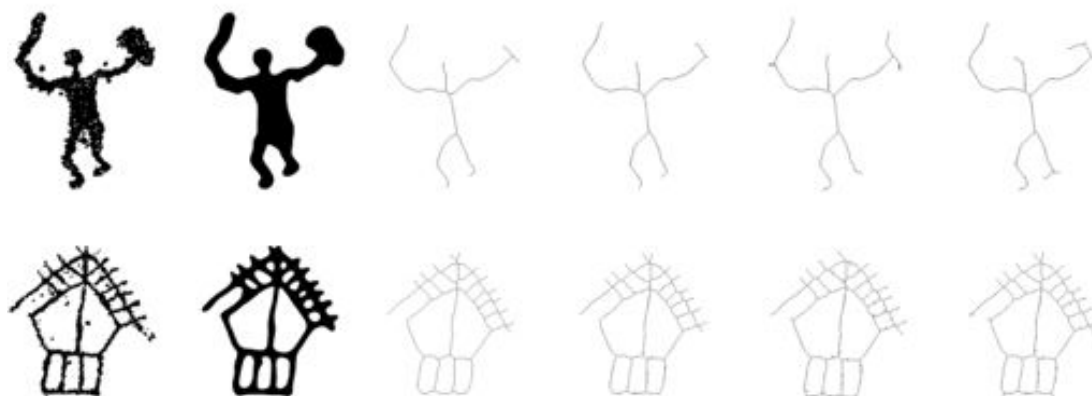


Figure 8.1: A human figure and a house, the pre-processed shapes and skeletonisation results.

THE shape description approach described in the previous chapter relies on robust skeletonisation of petroglyph shapes (see Figure 8.1). Skeletonisation and boundary-based descriptors rely on closed boundaries of a shape. In this chapter, we investigate both: a) the preprocessing of petroglyph shapes to close boundaries and b) the subsequent skeletonisation.

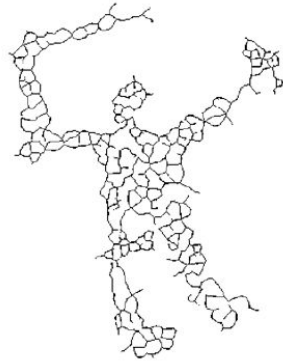
The petroglyph tracings we use (see Subsection 2.6.2) pose novel challenges to skeletonisation due to their complex topology and structure. They may be incomplete due to partial abrasion of the rock surface. Since the petroglyphs are made of individual peck marks they exhibit a complex boundary as well as numerous holes in their interior (see Figure 8.2(a)). Additionally, complex figures may consist of several disconnected parts. Petroglyph shapes can either show filled bodies or just the silhouette of a figure depending on their artistic style. Finally over the years figures have been pecked on top of each other which results in merged shapes.

Existing skeletonisation algorithms are not directly applicable to this type of material and yield poor skeletons as shown in Figure 8.2(b). One reason for the poor performance is that existing methods are usually developed on perfectly segmented

8. SKELETONISATION OF PETROGLYPH SHAPES



(a) Unprocessed shape from digitized petroglyph tracing



(b) Skeleton of unprocessed petroglyph shape



(c) Skeleton of pre-processed petroglyph shape

Figure 8.2: Comparison of the original petroglyph tracing (a) and its skeleton (b). The petroglyph shape pre-processed with the proposed method results in a significantly simplified skeleton that still preserves the important parts of the original shape (c).

shapes with continuous contour and continuously filled regions originating, for example, from public datasets such as MPEG-7 Core Experiment CE-Shape-1 and Kimia-99 [BLL07, BL07, YBYZ09, SBYL13]. Thus most methods do not fulfill the requirements of noisy real-world data such as that employed in this work. Aside from different applications (e.g. shape retrieval) enabled by robust skeletonisation, the employed petroglyph shapes pose a challenging testbed for the further development of skeletonisation algorithms.

Our contribution is a comprehensive investigation of skeletonisation in a new application domain that provides noisy real-world data and has rarely been explored so far [TTMI06]. We study the applicability of existing skeletonisation methods and evaluate their strengths and weaknesses. We identify two major requirements for robust skeletonisation: (i) pre-processing of the shapes is necessary to smooth the contour and fill holes in the interior; (ii) existing methods require adaptations to handle the complex structure of the shapes. We propose several improvements of recent skeletonisation methods to make them applicable to the investigated material. Finally, we come up with a novel shape pre-processing method that improves the shapes by smoothing and hole filling. The goal of the pre-processing is to reduce the complexity of the shapes and thereby to facilitate the subsequent skeletonisation. Our method balances the degree of smoothing and hole filling by an adaptive threshold to yield a suitable tradeoff between preserving the topology and removing noise. An evaluation on our dataset shows that pre-processing is a crucial step for robust skeletonisation of noisy and complex shapes like petroglyph tracings (see Figure 8.2). Furthermore the proposed improvements of the skeletonisation algorithms yield more accurate and complete skeletons.

The chapter is structured as follows: In Section 8.1 we review related work on skeletonisation and skeleton pruning and identify suitable algorithms for our task. Subsection 8.2.1 presents our real-world material and its characteristics. We describe our pre-processing approach and the improvements of skeletonisation algorithms in Section 8.2. Experimental setup and results are presented in Sections 8.3 and 8.4. Finally we draw conclusions in Section 8.5.

8.1 Related work

In this section we review skeletonisation and skeleton pruning algorithms, analyze their properties and identify suitable methods for our task. In the literature the usage of terminology for skeletonisation is highly ambiguous. Skeletonisation, thinning, medial axis transform, distance transform as methodologies and skeleton, medial axis or medial line as their results are used inconsistently [Baj06]. According to Arcelli and Bajaj [AB96] algorithms for skeleton computation in *discrete space* can generally be partitioned into two categories: Methods that perform *skeletonisation by medial axis transform* produce skeletons following Blum’s definition of the medial axis [Blu67] and techniques employing *skeletonisation by thinning* derive a thin version of a shape [Din55, KCRU58]. A third category of approaches performs the medial axis transform to polygonal shapes in *continuous space* [Mon69, Kir79, OI92]. Additionally, there is a

8. SKELETONISATION OF PETROGLYPH SHAPES

group of more recent skeletonisation algorithms that utilize physics-based modeling of the shapes [Par11, KC09] for which we suggest a fourth category.

All skeletonisation methods are sensitive to boundary noise, i.e. small perturbations of a shape may have large influence on the skeleton (see Figure 8.2). To overcome this problem some form of regularization is required [SB98]. This regularization process is generally referred to as “*skeleton pruning*”. Shaked and Bruckstein observe that pruning is an essential part of skeletonisation algorithms and most recent developments combine skeletonisation and skeleton pruning in one algorithm. Skeleton pruning methods can be consolidated in two major categories: The first covers the pruning of skeleton branches based on a significance value calculated for every single skeleton point. This results in a shortening of all skeleton branches. The second class of skeleton pruning algorithms calculate a significance measure for each branch. Based on its significance value a branch is either removed completely or remains in the skeleton [LWZH13].

8.1.1 Point-based pruning approaches

Montanari first develops a form of regularization to detect the most important skeleton branches [Mon68]. He proposes the use of a threshold for Blum’s “*propagation velocity of the wavefront*”. Blum and Nagel extend this idea and propose a boundary/axis weight for the regularization of unwanted branches caused by boundary perturbations [BN78]. They state, however, that boundary perturbations are not always unwanted distortions but might actually be important features of a shape and therefore pruning should be carried out with great care. Ho and Dyer propose the computation of the relative prominence of a skeleton point by using geometric relations between the maximum generating disk at the point and the contour of the shape [HD84]. Ogniewicz and Ilg compare several other regularization methods for skeleton points and propose the generation of a skeleton pyramid for further pruning [OI92]. Telea and van Wijk introduce a skeletonisation algorithm based on a fast marching level set method (Augmented Fast Marching Method, AFMM) [TvW02]. For every skeletal point they determine the length of the boundary segment it originates from and prune skeleton points using a single threshold. Howe applies the work of Telea and van Wijk to handwriting recognition using the contour length as saliency measure [How04]. Shen et al. compare this and other pixel-based significance measures and introduce a new significance measure for skeleton pruning by calculating the bending potential ratio (BPR) of the contour segment generated by the two points of the maximum inscribed disc that are tangent to the boundary [SBH⁺11]. Telea further improves AFMM by a different saliency metric, and proposes skeletonisation for feature preserving shape smoothing [Tel12].

8.1.2 Branch-based pruning approaches

The methods summarized in Subsection 8.1.1 all compute a significance value for each single point of the skeleton. A thresholding of this value leads to a shortening of the branches. Branch-based methods, in contrast, avoid the shortening of branches and instead use a significance value to remove or retain entire branches. Bai et al. [BLL07]

propose a novel method for skeleton pruning based on Discrete Curve Evolution (DCE) introduced by Latecki and Lakamper [LL99]. They determine the contour points of a shape that have maximum curvature and delete all skeleton branches that do not end at one of these points. This approach inspired numerous other state-of-the-art skeleton pruning algorithms. Bai and Latecki further improve DCE by removing the necessity of prior knowledge about the shape [BL07]. They compute the DCE-skeleton with a fixed parameter (50 vertices) and subsequently add a reconstruction step, which removes skeleton branches with low contribution to the original shape. Yang et al. use the same methodology and extend the reconstruction algorithm to increase speed and to enable the computation of skeletons from shapes with holes [YBYZ09]. Shen et al. introduce a normalization factor in the reconstruction step that quantifies the tradeoff between the simplicity of their skeletons and the reconstruction error of the shapes [SBYL13]. Liu et al. extract the Generalized Voronoi Skeleton of a shape and then apply DCE to perform a first pruning of the obtained skeleton [LWH⁺12]. Subsequently, they further prune by balancing the visual contribution and the reconstruction contribution of each skeleton branch. Liu et al. further devise a skeleton pruning approach that fuses the information of several different branch significance measures [LWZH13]. Recently, Krinidis and Krinidis proposed a new skeletonisation approach that smoothes the polygonal approximation of a shape iteratively [KK13]. In each iteration they determine the most important polygon vertices from the angles of their incoming edges and prune the skeletons by deleting those branches that connect less important nodes.

8.1.3 Comparison of algorithms

In the following, we specify a number of criteria for the comparison of the skeletonisation and pruning techniques presented above.

- *Robustness against remaining insignificant branches:* Insignificant branches are branches that do not contribute essentially to the original shape and thus should be avoided or pruned.
- *Robustness against deletion of significant branches:* A significant branch has an essential contribution to the figure's shape and thus should remain in the skeleton. A deletion would significantly change the structure of the skeleton.
- *Robustness against branch shortening:* Branch shortening occurs when insignificant as well as significant skeleton branches are shortened likewise. This bears the risk of changing the structure of the skeleton.
- *Rotation and scale invariance:* The skeleton of a differently scaled and rotated shape should be equivalent.
- *Number of parameters:* A large number of parameters increases the dependence of an algorithm on user input but at the same time gives more control. We prefer algorithms with a low number of parameters with adequate sensitivity to parameter changes.



Figure 8.3: Some shapes from example classes (e.g. antropomorph, bird, cross, deer, etc.) from our entire dataset of more than 1100 shapes classified into two different archeological typologies.

- *No prior knowledge about shape needed:* Parameters such as the number of end-points, or absolute values that depend on the size and complexity of the shape, require *a priori* knowledge and should be avoided.

Table 8.1 compares the methods presented above with respect to the specified criteria. The comparison shows that all of the reviewed algorithms differ from each other on at least one criterion. Some of them show major deficiencies such as the deletion of significant skeleton branches, rotation and scale variance and the need for prior knowledge about the shapes. For our experiments on petroglyph skeletonisation we select a subset of the presented algorithms. We perform the selection of methods based on the following considerations: (i) both branch-based and point-based pruning methods should be selected to increase the heterogeneity of evaluated approaches; (ii) all algorithms should be robust to scale and rotation, as the investigated petroglyphs may be rotated and scaled arbitrarily; (iii) the number of input parameters should be low. Based on these considerations, we select the *BPR*-algorithm of Shen et al. [SBH⁺11], the *DCE*-algorithm of Bai et al. [BLL07], and the *SPT*-algorithm of Shen et al. [SBYL13]. Furthermore, we add simple morphological *thinning* as an additional method to investigate a skeletonisation method that does not apply any pruning.

8.2 Approach

The characteristics of petroglyph shapes impede skeletonisation which leads to unsatisfactory results (see for example Figure 8.2). For robust skeletonisation a pre-processing

Table 8.1: Comparison of recent skeletonisation methods with respect to the identified criteria. Note that for point-based pruning approaches the first two criteria do not apply because they do not distinguish between significant and insignificant branches.

	Robustness against remaining insignificant branches	Robustness against deletion of significant branches	Robustness against branch shortening	Rotation and scale invariance	Number of parameters	No prior knowledge about shape needed
Point-based pruning						
Chord residual + Skeleton pyramid [OI92]	-	-	no	yes	3	no
AFMM [TvW02]	-	-	no	no	1	no
Boundary length [How04]	-	-	no	yes	1	no
Bending potential ratio [SBH ⁺ 11]	-	-	no	yes	1	yes
Saliency metric [Tel12]	-	-	no	yes	1	yes
Branch-based pruning						
DCE-skeleton [BLL07]	no	no	yes	yes	1	no
Discrete skeleton evolution [BL07]	yes	yes	yes	yes	1	yes
Quick stable skeletons [YBYZ09]	yes	yes	yes	yes	1	yes
Tradeoff reconstruction error / skeleton simplicity [SBYL13]	yes	yes	yes	yes	2	yes
Visual contribution / reconstruction con- tribution [LWH ⁺ 12]	yes	yes	yes	yes	3	no
Information fusion [LWZH13]	yes	yes	yes	yes	1-2	yes/no
Empiric mode decomposition [KK13]	yes	yes	yes	yes	1	yes

8. SKELETONISATION OF PETROGLYPH SHAPES

of the shapes is necessary as well as improvements of skeletonisation techniques. Subsequent to the description of our material, we present a fully automated shape pre-processing method and propose a number of improvements for the selected skeletonisation algorithms to make them applicable to petroglyph shapes.

8.2.1 Investigated material

For the skeletonisation experiments in this chapter, we use the full dataset as described in Subsection 2.6.2. Figure 8.3 shows a small subset of the dataset. Earlier experiments of Takaki et al. on Central Asian petroglyphs showed that skeletonisation is a useful abstraction of shapes [TTMI06]. Thus it enables higher level applications such as similarity search and automated shape classification, which is our ultimate goal. Petroglyphs pose a challenge for skeletonisation as they are made of single peck marks and thus have neither a continuous contour nor continuously filled regions. Figure 8.3 shows that the shapes have highly varying complexity, contain numerous holes due to incompletely pecked areas, often contain very fine structures (horns of deer, feathers of birds, etc.) and have disconnected parts.

8.2.2 Adaptive shape pre-processing

The major challenge of pre-processing is to find a solution that forms a good tradeoff across all different types of petroglyphs in our dataset. We define a minimum set of requirements that a suitable pre-processing needs to fulfill to enable more robust skeletonisation: (i) close small holes, (ii) smooth the contour, (iii) connect nearby parts, and (iv) avoid the unwanted decomposition of the shapes. We design a method that automatically improves the shapes according to these requirements and that terminates autonomously when a good tradeoff amongst the four aims is achieved. The intuition behind our approach is to combine different structural and morphological filters in a way that they join their strengths. An overview of the method is shown in Figure 8.4.

Initially we resize and pad all shapes to normalize the inputs. Next we apply a median filter with size s_{med} to the input. Median filtering removes small holes in foreground and background (salt and pepper noise) and at the same time slightly smooths the contour. Apart from this, however, the median filter may generate artifacts by disconnecting weakly connected blobs. To compensate for these artifacts, we apply an area opening and closing originally proposed by Vincent for grayscale images [Vin94]¹ as well as a dilation operation. We use an area size of t_{aoc} pixels as threshold for area opening and closing, and combine it with a dilation by a disc with a radius r_{dil} to reconnect disjoint parts. We iterate these steps with increasing median filter size until a stopping criterion is met.

The stopping criterion requires a robust indicator function that is suitable for the differently complex shapes in the dataset. We evaluate different indicator functions such as solidity and circularity of the shape, the number and size of foreground and

¹Note that the proposed area opening/closing is fundamentally different from a morphological opening/closing as it does not employ a structuring element.

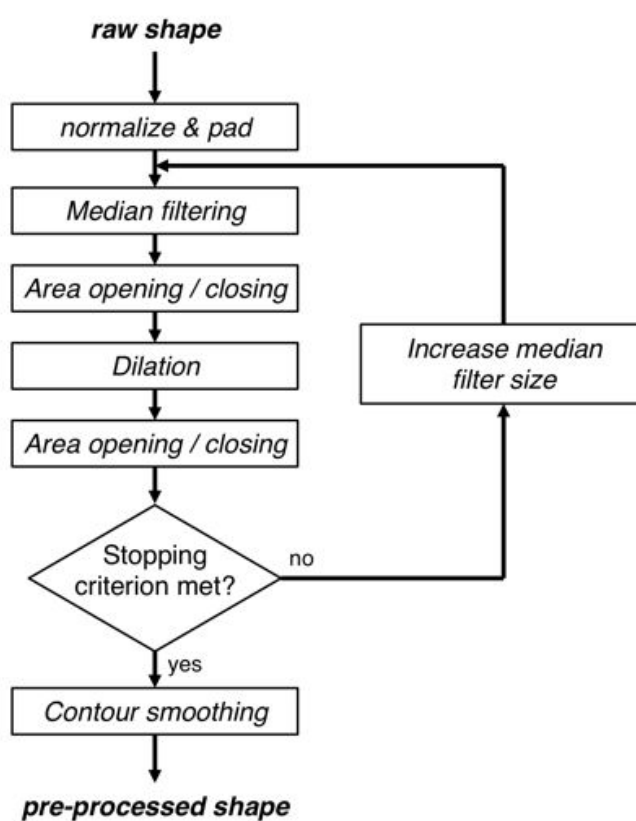


Figure 8.4: The workflow of the automated shape pre-processing method. We input a raw binary shape and apply a median filter and binary morphology.

8. SKELETONISATION OF PETROGLYPH SHAPES

background blobs, and the number of endpoints in the thinning skeleton. Our preliminary experiments show that the most robust criterion is a combination of the number of foreground blobs and the number of background blobs. If the number of both do not change over a certain number of iterations n_{it} , we terminate the pre-processing. The assumption behind this stopping criterion is that if the number of foreground and background blobs remains constant for some time, the figure is likely to be in a robust state where the influence of noise is low.

After the iterative pre-processing has terminated, we smooth the contour by a convolution filter with a Hann window of size s_{conv} . We smooth the x- and y-coordinates of the shapes' contour points separately which removes contour perturbations and thus reduces the likelihood to get spurious (insignificant) skeleton branches in subsequent skeletonisation.

Figure 8.5 illustrates multiple iterations of the proposed pre-processing method for a given shape. The numbers of foreground and background blobs decrease rapidly during the initial iterations as holes in the shape are removed and closely spaced shape parts are merged. After four iterations the shape reaches a robust state which is indicated by the fact that the numbers of foreground and background blobs do not change for a certain number of iterations, i.e. the two indicator functions reach a plateau. We observe further plateaus (e.g. starting at iteration 7, 12, and 15). As iterations proceed the plateaus usually get longer. The stopping criterion finally decides after which plateau length to stop by parameter n_{it} . For the given example $n_{it} = 3$ yields satisfactory results. After the first plateau the number of foreground blobs increases again and the shape starts to decompose.

8.2.3 Improvements of existing skeletonisation methods

We take the automatically pre-processed shapes as basis for the computation of the skeletons with the chosen algorithms (thinning, BPR, DCE, and SPT). In our preliminary experiments we observe that two of these algorithms (DCE and SPT) suffer from severe problems (see Figure 8.6(a), left column): The original DCE algorithm [BLL07] is not able to compute the skeleton of multiple blobs in one image, it does not preserve the topology of the shape (loses inner connections) and sometimes generates incomplete skeletons. We propose improvements to compensate for these shortcomings. First, we apply skeleton computation in all blobs of the input instead of on the first blob only (see Figure 8.6(a), 1st row). To preserve the inner connections, we extend the contour tracing of DCE in a way that all inner contours are traced (see Figure 8.6(a), 2nd row). The original DCE-algorithm employs either only the concave or only the convex contour points (depending on which set of points is larger) for skeletonisation and misses significant shape parts when the concave are selected (see Figure 8.6(a), 3rd row). We remove the selection criterion and employ all (convex and concave) points for skeletonisation. As a result we obtain complete skeletons even for complex shapes. The effects of our improvements are shown in Figure 8.6(a), right column. The SPT algorithm [SBYL13] utilizes DCE for a coarse pruning of the shape and thus benefits from our improvements as well. A second pruning step in SPT refines the skeleton by deleting branches whose

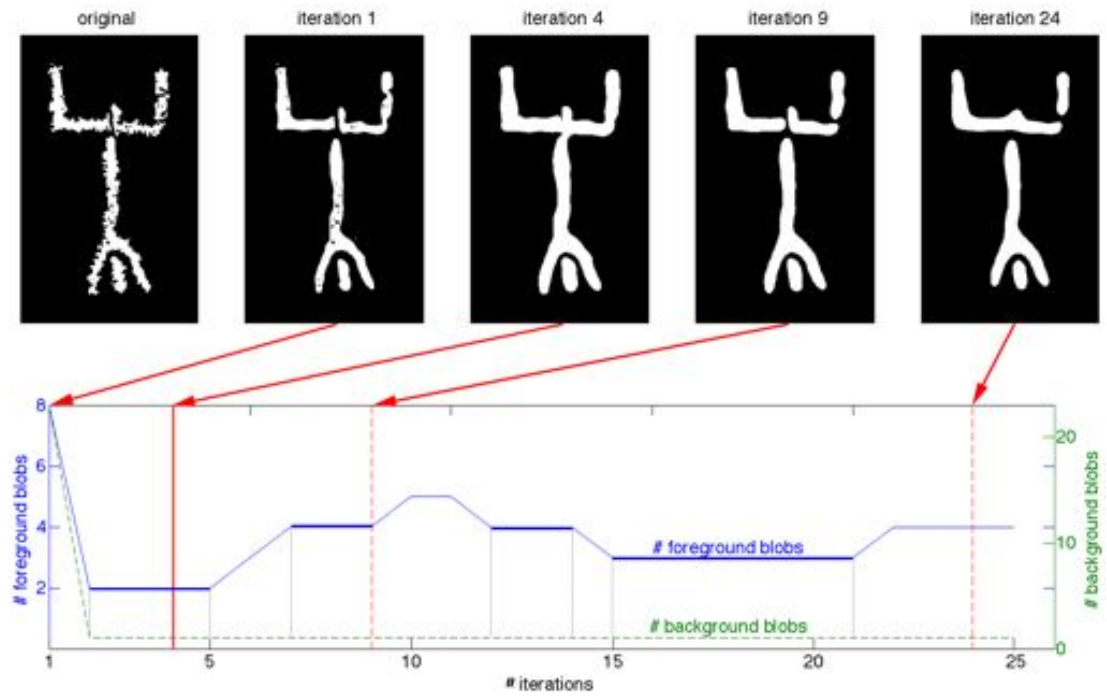
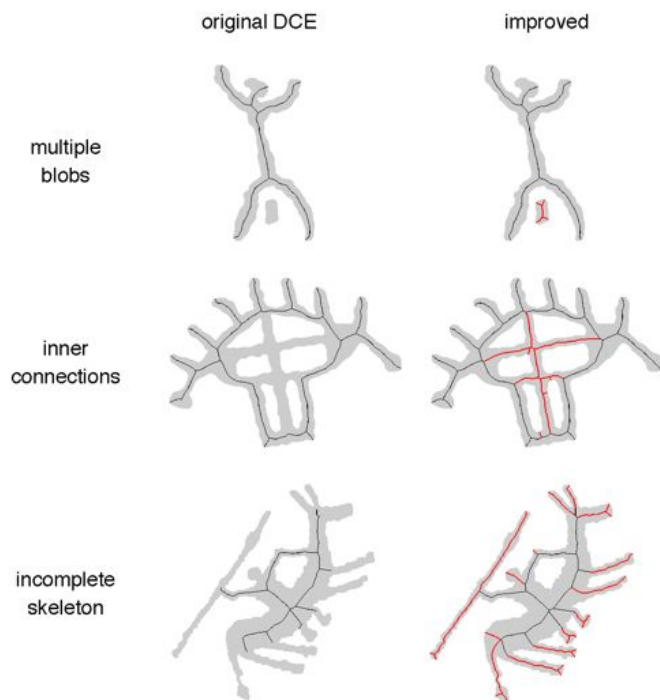
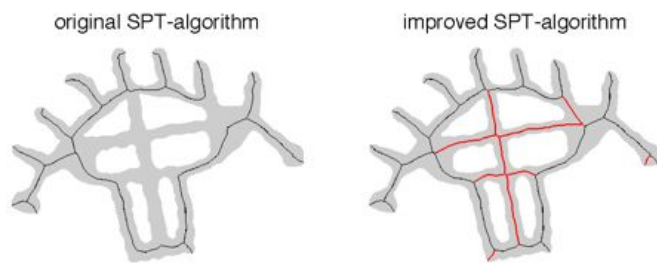


Figure 8.5: Sequence of the automated pre-processing of a shape with the proposed method. The numbers of foreground and background blobs decrease rapidly. At iteration 2, a stable minimum of number of blobs is reached and the stopping criterion aborts the loop at iteration 3. We observe that further iterations decompose the shape starting at iteration 6.

8. SKELETONISATION OF PETROGLYPH SHAPES



(a) Original (left column) and improved DCE-algorithm (right column)



(b) Original (left) and improved SPT-algorithm (right)

Figure 8.6: Comparison of skeletons of the original DCE and SPT algorithms (left column) with our improved versions (right column). Our improvements eliminate all shortcomings.

reconstruction contribution to the shape is below a certain threshold. In the original reconstruction step the skeleton path tracing starts at one seed point and traces the skeleton to every endpoint. Thereby inner connections and partly outer connections are missed and thus removed from the skeleton (see Figure 8.6(b), left). We improve the skeleton path tracing by setting several seed points at the skeleton branch points. As a result all inner connections are preserved in the final skeleton (see Figure 8.6(b), right).

8.3 Experimental setup

From the annotated material described in Subsection 2.6.2, we derive a dataset of 1138 petroglyph shapes to carry out our experiments with. In the following, we describe the selection of the parameters for our adaptive shape pre-processing method and the investigated skeletonisation methods and present our evaluation criteria.

8.3.1 Selection of parameters

For the selection of suitable values for the parameters defined in Subsection 8.2.2, we evaluate the pre-processing method on a reduced dataset consisting of 150 representative shapes from the entire dataset. In the absence of a ground truth of pre-processed shapes, we face difficulties in the selection of suitable parameter values. Hence we choose a heuristic approach to estimate robust parameters for the proposed method.

To determine the area opening and closing threshold t_{aoc} [Vin94], we examine the histogram of all blob sizes in the reduced dataset. The histogram is strongly skewed to the left and exhibits a peak between blob sizes of 5 and 20 pixels. Blobs of this size mostly represent noise (e.g. small holes) that usually does not contribute to a figure’s shape significantly. With a threshold t_{aoc} between 5 and 20 pixels about 20% of all blobs are removed and a large portion of noise is filtered out. For the remaining experiments with the entire dataset, we set t_{aoc} to 10 pixels. For the size of the dilation disk r_{dil} we select a rather low value of 3 pixels. This facilitates the reconnection of nearby parts and minimizes the amount of region growing so that the likelihood of merging unrelated shape parts is reduced. From the reduced dataset we observe that typical plateau lengths of the indicator functions are between 2 and 5 iterations. A value of $n_{it} = 3$ iterations yields a good tradeoff between the sufficient smoothing of perturbed shapes, and the decomposition of thin and sparsely connected figures. To estimate s_{conv} , we apply thinning to the shapes in the reduced dataset and count the number of endpoints. We repeat this process with increasing values of s_{conv} . For filter sizes s_{conv} between 30 and 70 contour points the number of the endpoints remains mostly constant. This shows that the parameter has low sensitivity. We set s_{conv} to 51 pixels for our experiments. The size of the median filter is a non-critical parameter as it is simply increased every iteration. We start with a minimum size of $s_{med} = 3$ pixels and increase it by 2 pixels every iteration.

Additional parameters have to be selected for the skeletonisation methods. For *DCE* [BLL07], we estimate the parameter for the number of vertices adaptively by counting the number of endpoints of the respective thinning skeleton. For the *BPR*-algorithm [SBH⁺11] and the *SPT*-algorithm [SBYL13], we take the parameter values as proposed by their authors. The computation of the *thinning* skeleton is parameter free as it is a simple morphological operation.

8. SKELETONISATION OF PETROGLYPH SHAPES

8.3.2 Evaluation

Due to the absence of ground truth shapes and skeletons, we define several perceptual evaluation measures that can easily be judged by a human observer. Subsequently, we evaluate our pre-processing method and the applied skeletonisation algorithms separately on the entire dataset.

For the evaluation of the pre-processing the following criteria are considered:

- All shape parts that are important for visual perception are preserved.
- Independent but closely spaced parts are not merged (e.g. legs or feathers).
- Small holes in the shape are closed, disjoint parts are reconnected and the contour is properly smoothed, i.e. the shape is likely to facilitate subsequent skeletonisation.

For the evaluation of the obtained skeletons we apply the following criteria:

- The skeleton preserves the full structure of the shape.
- It exhibits branches for all important parts of the shape.
- It does not have remaining spurious branches.

8.4 Results

First, we evaluate the effect of the proposed pre-processing method quantitatively. The intuition behind this evaluation is to obtain an indication of how effectively the pre-processing simplifies the shapes (and thus improves them for subsequent skeletonisation). To assess this ability, we utilize a simple observation, namely that typical petroglyph skeletons should usually not exhibit more than 20 endpoints. This can be derived by looking at the shapes in Figure 8.3. A consequence of this observation is that skeletons with more than 20 endpoints are likely to contain spurious branches. We use the number of endpoints as a simple heuristic to evaluate the performance of the pre-processing method. For evaluation we count the number of endpoints of the thinning skeletons before and after pre-processing. Figure 8.7 shows the distribution of the number of endpoints before and after pre-processing. Pre-processing strongly reduces the number of endpoints in the resulting skeletons and maps them to a reasonable range. Thus this evaluation provides a first indication that the proposed pre-processing is beneficial for subsequent skeletonisation.

As the employed shape pre-processing operations change the original shapes, for a more comprehensive (qualitative) evaluation of the pre-processing methods a visual inspection of the generated shapes and skeletons is necessary. For this purpose, we manually judge each automatically pre-processed shape according to the criteria presented in Subsection 8.3.2. The evaluation results for the proposed pre-processing method on the entire dataset are shown in Table 8.2.

The pre-processing method performs well for 79.8% of all figures. Figure 8.8(a) shows examples where unimportant holes in the shapes are removed successfully and

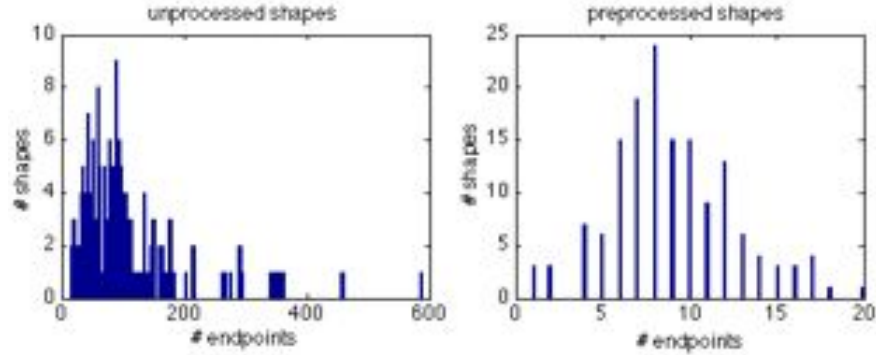


Figure 8.7: Comparison of histograms of the number of endpoints of the thinning skeletons of unprocessed shapes (left) with the pre-processed shapes (right). The number of endpoints significantly decreases and is mapped to a reasonable range.

Table 8.2: Results of the quantitative evaluation of the proposed pre-processing method on the entire dataset.

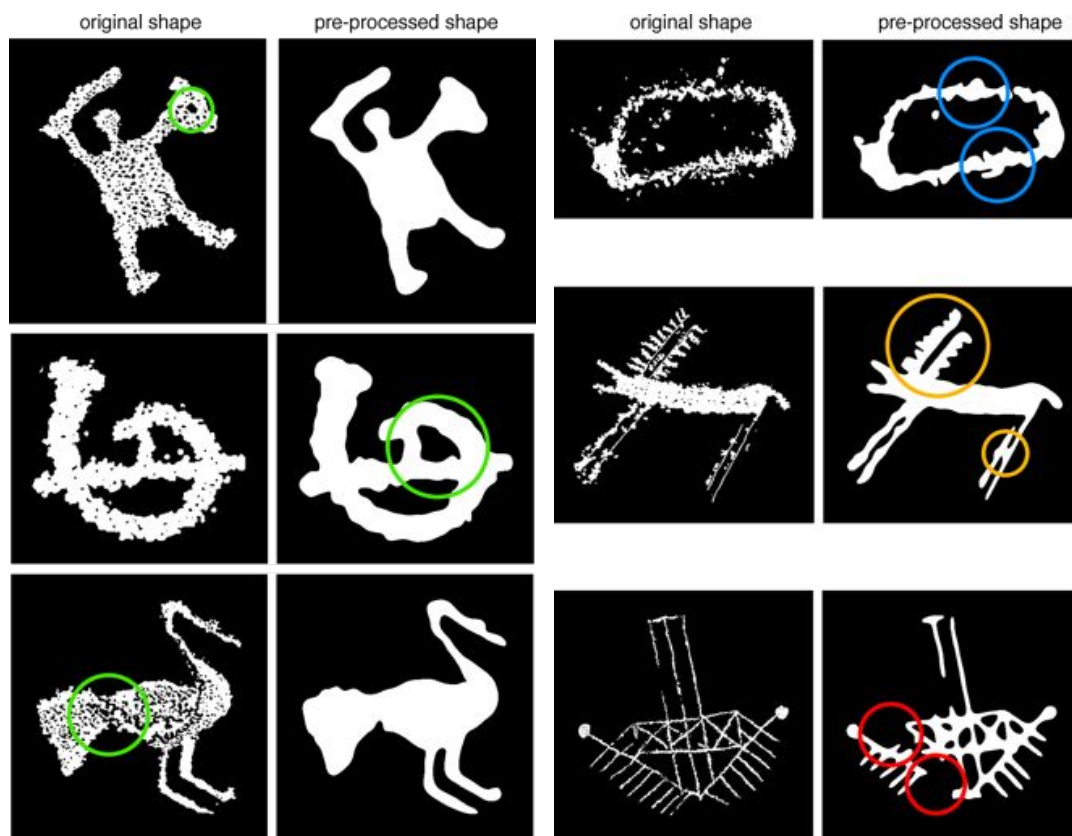
Not smooth enough	Merged parts	Lost details	\sum errors
9.4%	6.9%	3.9%	20.2%

the contours are properly smoothed. For the remaining 20.2% of the shapes we identify different types of errors (see Figure 8.8(b) and Table 8.2). 9.4% of all shapes are not smoothed sufficiently and their contours remain perturbed, which impedes skeletonisation. For 6.9% of the shapes closely spaced parts are merged although it would be important for skeletonisation that they remain separated. Only 3.9% of the shapes are smoothed too strongly and thus important shape parts are lost. The proposed method thus forms a tradeoff between the filtering of noise and removal of unimportant shape parts and the smoothing of important shape parts.

Similarly to the evaluation of the pre-processing, we assessed the quality of the skeletonisation methods according to the criteria presented in Subsection 8.3.2. The performance evaluation of the skeletonisation methods on the pre-processed shapes for the entire dataset is shown in Table 8.3. On average (over all skeletonisation algorithms) we obtained satisfactory results for the majority of our shapes. Figure 8.9 provides examples of the obtained skeletons and demonstrates that the skeletons are accurate to a high degree. We comprehensively investigated different types of errors that arise during skeletonisation and observed that all skeletonisation algorithms have deficiencies in certain situations, but none of them completely fails. See Figure 8.10 for examples of improper skeletonisation.

The skeletons pruned with DCE are in most cases complete (96.3%), but the algorithm additionally produces a lot of spurious branches (47.7%). This is related to the fact that the algorithm requires prior knowledge about the shapes (the number of DCE

8. SKELETONISATION OF PETROGLYPH SHAPES



(a) Holes are filled, contours are smoothed, and (b) Figures are not smoothed sufficiently (top), the overall shape is preserved (middle). shape parts are merged (middle), and details are lost (bottom).

Figure 8.8: Comparison of the results of satisfactory (a) and weak (b) pre-processing of petroglyph shapes.

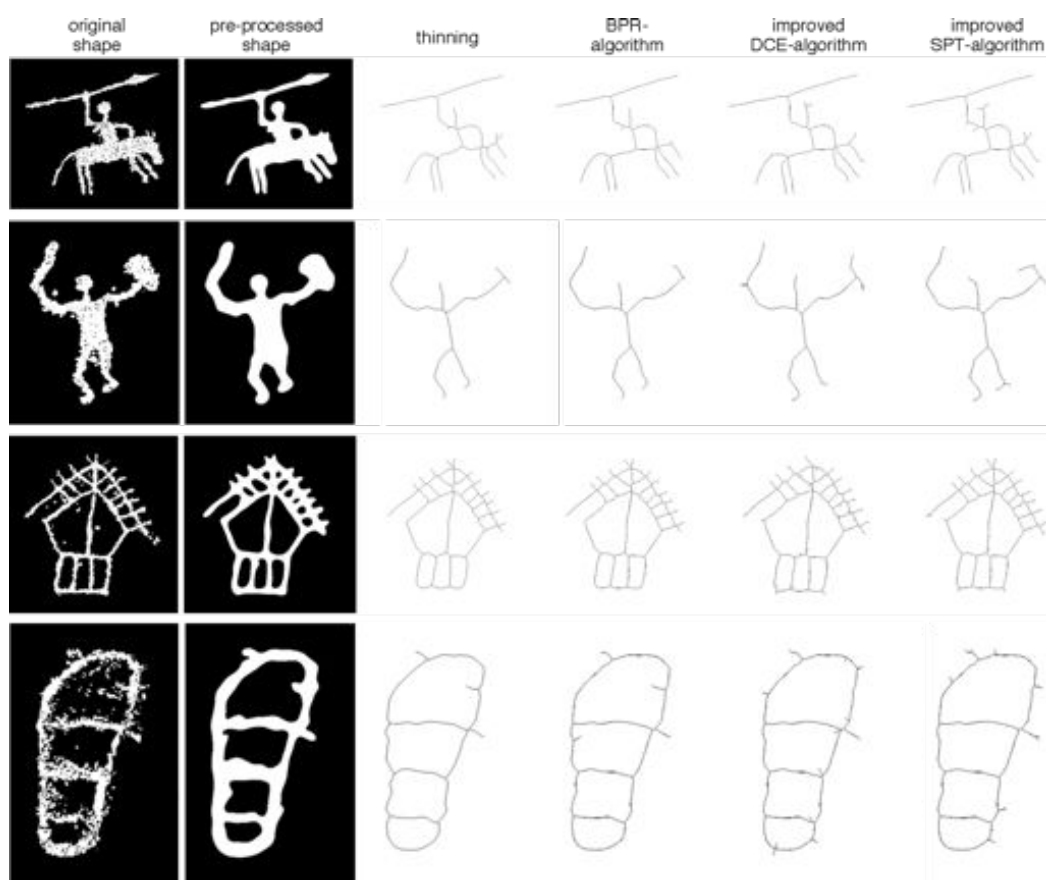


Figure 8.9: Examples of successful skeletonisation with all selected and improved algorithms based on shape pre-processing with the proposed method. Note that for DCE and SPT a few spurious branches remain in contrast to thinning and BPR.

8. SKELETONISATION OF PETROGLYPH SHAPES

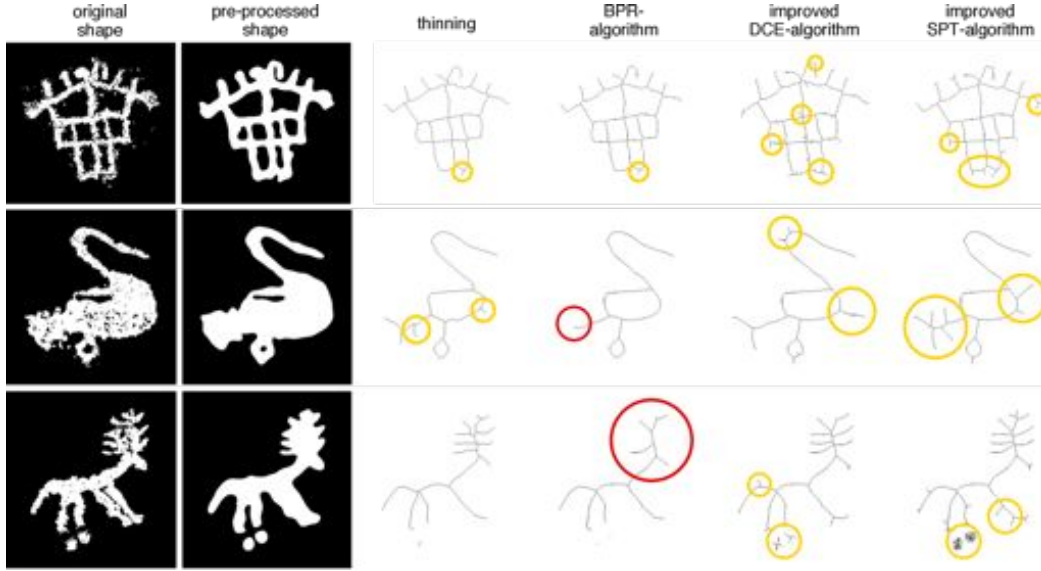


Figure 8.10: Examples of partly erroneous skeletons obtained with the selected algorithms. In all depicted examples spurious branches exist (yellow circles) and for some shapes important parts are deleted (2nd and 3rd row, red circles).

vertices). If this number is set too low, significant branches are deleted even before spurious branches (3.7% of all shapes). If the number is set too high, a lot of spurious branches remain (47.7%). We set this parameter adaptively (depending on the number of endpoints of the corresponding thinning skeleton) since a unique number that is suitable for all shapes does not exist.

SPT builds upon DCE and performs similarly. Although it produces even more spurious branches (65.8% of all shapes), it deletes spurious branches *before* important ones. Thus for only 1.4% of all shapes important parts are lost. An advantage of SPT over DCE is that it does not require *a priori* information about the shapes.

The BPR algorithm outperforms DCE and SPT and produces satisfactory skeletons for 86.9% of all shapes. BPR generates much fewer spurious branches than DCE and SPT (only 6.7% of all shapes). The BPR pruning is, however, in some situations too strong and thus important branches are removed in 6.4% of all shapes (see for example tail of the bird in Figure 8.10, 2nd row).

A simple thinning results in notably good skeletons for 83.5% of all shapes which is nearly as good as the performance of the more sophisticated BPR algorithm. Since petroglyph shapes often resemble stick-like figures, they can be well modeled by the thinning algorithm. Additionally, the contour smoothing in the pre-processing avoids the generation of spurious branches by thinning (in 87.5% of all cases). The results obtained for thinning show that a proper pre-processing can replace an additional skeleton pruning. In only 4.0% of all cases important branches are missed by thinning.

Table 8.3: Evaluation of the four selected skeletonisation algorithms on the entire dataset. BPR and thinning perform best on the pre-processed shapes.

Problem \succ Algorithm \vee	Spurious branches	Lost parts	\sum errors
Thinning	12.5%	4.0%	16.5%
BPR [SBH ⁺ 11]	6.7%	6.4%	13.1%
DCE [BLL07]	47.7%	3.7%	51.4%
SPT [SBYL13]	65.8%	1.4%	67.2%

8.5 Conclusion

In this chapter, we presented a study on skeletonisation of petroglyph shapes. We studied the applicability of existing skeletonisation methods and evaluated their strengths and weaknesses. Existing skeletonisation methods were developed and evaluated mainly on ideal shapes and are thus not directly applicable to our real-world data. Therefore we improved several skeletonisation algorithms to compensate for their shortcomings that became apparent. Additionally, we proposed an adaptive shape pre-processing method that enables the computation of robust skeletons for the complex and diverse shapes under investigation. We performed a large-scale experiment and showed that proper pre-processing is crucial for the skeletonisation of petroglyph shapes. Experiments on skeletonisation showed that pre-processing in combination with a simple thinning yields a good tradeoff for robust skeletonisation, whereas more sophisticated skeletonisation techniques either generate more spurious branches (DCE, SPT) or delete important ones (BPR). Our experiments clearly demonstrated that the presented pre-processing method and the proposed improvements of recent skeletonisation methods solve the additional challenges introduced by our complex and noisy real-world shape data for more than 86% of all investigated shapes.

8. SKELETONISATION OF PETROGLYPH SHAPES

Automated Classification of Petroglyph Shapes



Figure 9.1: A query petroglyph (left) and the most similar petroglyphs determined by our system.

IN this chapter we investigate three questions. First, whether our approach lined out in Chapter 7 generalizes to a larger dataset; second, how the employed descriptors and descriptor combinations perform in detail as well as in comparison to popular state-of-the-art methods for image classification (deep learning) and third, whether the results can be incorporated into the workflow of petroglyph documentation.

In Section 9.1, we describe our automated classification approach and introduce related work not mentioned in previous chapters. In Subsection 9.2.1, we evaluate the petroglyph descriptor combination (see Figure 9.1) proposed in Chapter 7 on a larger dataset. Subsection 9.2.2 presents details on the performance of single descriptors and compares to a state-of-the-art visual deep learning framework. Section 9.3 presents some thoughts about the value-in-use of our tools while section 9.4 presents our conclusion.

9.1 Approach and Related Work

For the experimental evaluation in this chapter, we use the approach proposed in Chapter 7. In this section we describe the most important points. For a detailed description of the approach and detailed evaluation results on a small petroglyph dataset please refer to Chapter 7.

Numerous surveys about shape analysis have been published, comprehensive and important in this field are the surveys by [Pav78], [Lon98], [ZL04] as well as by [YKR08]. Classifications and taxonomies of shape descriptors have been proposed in the mentioned surveys in different variants. A widely used common denominator is the distinction between contour-based and region-based descriptors. Our intuition lined out in Chapter 7 was that the petroglyph shapes in our dataset are perceptually close to skeletons. Hence, we added a third category of skeleton-based approaches and categorized the described approaches in region-based, contour-based and skeleton-based approaches that are relevant for our work (see Section 7.1).

In Chapter 7, we combined the strengths of contour-based as well as region-based shape descriptors with the strengths of undirected skeletal graphs for the petroglyph classification problem. The methods we combined are a) Shape Context (SC) [BMP02], which has been used by Deufemia et al. [DP13] for petroglyph classification, b) Inner Distance Shape Context (IDSC) [LJ07], c) Generalised Hough Transform (GHT) proposed by Zhu et al. [ZWKL11] for petroglyph classification.¹ as well as d) Our own petroglyph descriptor which makes use of topological features of the undirected skeletal graphs of petroglyphs in order to create a feature vector for each petroglyph by graph embedding (GE).² In order to be able to use skeletonisation and contour-based descriptors, we needed to pre-process our material to achieve continuous boundaries. This pre-processing is described in Chapter 8. The resulting shape images are the input for SC, IDSC as well as for the skeletonisation for which we use the method of [BLL07].³ We normalize the width of the input petroglyph to 500px and prune the skeletons by joining nodes which have a spatial distance smaller than a given threshold. GHT is computed on the original petroglyph images.⁴

In this chapter, we build on the experiments in Chapter 7 where it turned out that a) GE and GED perform comparably well (with way lower computational demand for GE) and b) the combination of region-based, contour-based and skeleton-based descriptors yields the best results.

We employ a larger dataset (see Subsection 9.2.1) and additionally investigate the region-based Generic Fourier Descriptor (GFD, [ZL02]) Furthermore, we use a state-of-the-art deep learning framework which we tune for petroglyph shape shape recognition (see Subsection 9.2.2).

¹For all three descriptors, implementations have been kindly provided by the respective authors.

²We utilize the MIT strategic engineering tools for network analysis kindly provided by [BdW12]: http://strategic.mit.edu/downloads.php?page=matlab_networks.

³We thank Bai et al. for providing the implementation.

⁴We use resolutions of 10x10px, 20x20px and 30x30px. We report on the best performing resolution 10x10px.

GFD has been employed by Deufemia and Paolino [DP14]. They use an enhanced version of Generic Fourier Descriptors (EGFD) to detect and classify petroglyphs from full scenes. In their experiments on 1215 petroglyphs in 53 scenes EGFD outperforms the descriptors GHT and SC. They do not use combinations of descriptors.

We use the popular deep learning framework Caffe by Jia et al. [JSD⁺14] twofold: First, we use the features extracted with the pre-trained imagenet network that is included in the distribution of Caffe¹ (CNN_IN). Second, we tune the network by re-training the last layer of the network with our petroglyph shape data (CNN_RT). The re-training of a convolutional neural network is a computationally demanding task. We used leave-one-out cross-validation (LOOCV) in Chapter 7. This poses a dilemma. On one hand, we need to present results comparable to Chapter 7 and consequently have to use LOOCV. On the other hand, LOOCV requires the training of a net for each training data sample, which is in our case (1255 samples) computationally not feasible. Hence, we split the evaluation in two parts. In both parts, we use a combination of several shape descriptors and a k-NN classifier.

The results we report on in Subsection 9.2.1 aim at evaluating the performance and value-in-use of the descriptor combination proposed in Chapter 7. Thus, we validate with LOOCV to preserve backwards comparability to Chapter 7.

The detailed results in Subsection 9.2.2 extend the evaluation with GFD and deep learning, including a re-trained convolutional neural network (CNN_RT). As it is computationally not feasible to train a net for each training data sample (1255 samples) we employ 5-fold cross-validation in this section and repeat selected experiments from Subsection 9.2.1 with this cross-validation method for comparability.

9.2 Experiments and Results

9.2.1 Descriptor Combination Previously Proposed

In this section we apply the approach outlined in Subsection 7.2.2 to the large dataset described in Subsection 2.6.2. We investigate whether the good result - an accuracy of 88% - which was obtained on a small evaluation dataset in Chapter 7, can be generalised to a larger dataset. We analyze the data separately according to the two typologies that we have employed.

9.2.1.1 Modified Sansoni Typology

Figure 9.2 shows the classification results for the 26 classes examined based on the modified Sansoni typology. The overall accuracy is 65%. We observe that classes 1,2,7 and 14,15,16 yield clearly better accuracies (around 80%) than the overall accuracy. Table 2.1 shows that classes 1,2,7 and 14 are heavily populated, but 15 and 16 are only averagely populated with 35 and 43 examples. It is possible that classes 15 and 16 are visually homogenous and especially well suited for our classifier. Figure 9.4 shows

¹<http://caffe.berkeleyvision.org/gathered/examples/imagenet.html>

9. AUTOMATED CLASSIFICATION OF PETROGLYPH SHAPES

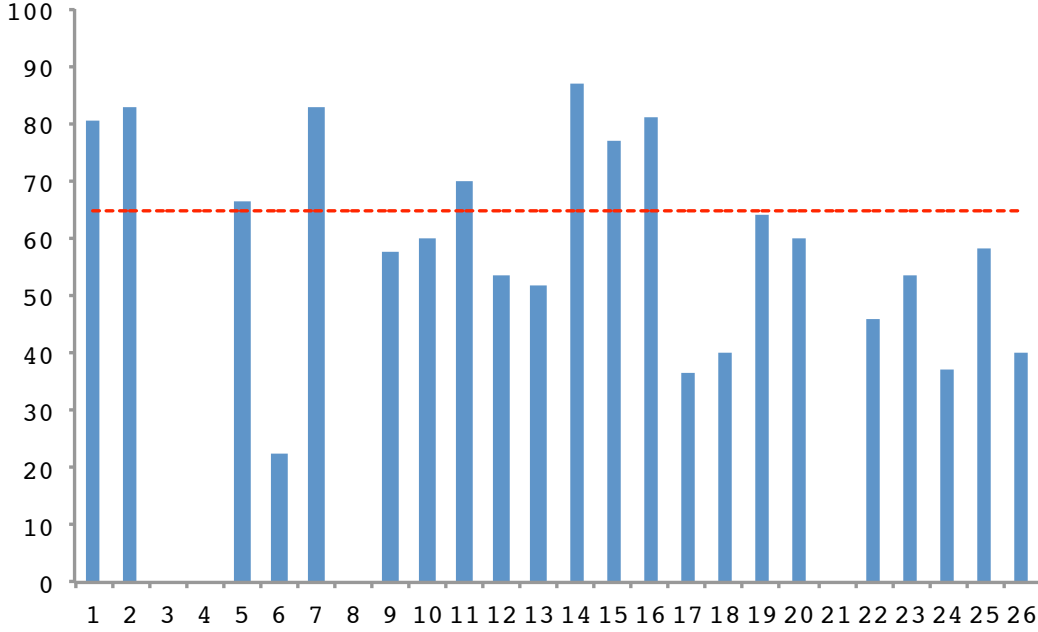


Figure 9.2: Classification accuracy of the classification results of the 26 classes based on the modified Sansoni typology. The red line shows the overall accuracy. Please refer to Table 2.1 for the class descriptions.

examples of class 16. We observe that the figures do, indeed, share a high degree of visual similarity. Note that our classifier is robust against changes of rotation and scale and so is relatively unaffected by differences of these types.

Furthermore we observe in Figure 9.2 that several classes (3, 4, 8 and 21) have an accuracy of 0%, i.e. not a single example of these classes has been correctly classified by our approach. Table 2.1 shows that classes 1 to 4 are variants of armed figures holding different weapons. The classes are unequally populated. Classes 1 and 2 contain 61 and 105 examples, classes 3 and 4 contain only 10 examples each. Figure 9.3 displays the confusion matrix of the classification results. Classification accuracy is good for classes 1 and 2 where there are sufficient examples for training. We observe that class 3 (which is shown in row 3) has the majority of its misclassifications being in class 2. Many examples of class 4 are also misclassified into class 2. We assume that the poor classification performance is caused by the small amount of training material for classes 3 and 4 and the fact that the visual differences between these armed figures are not sufficient for our descriptors. Class 8 contains 21 “Unarmed anthropomorph figures - not elsewhere classified”. The majority of these figures is misclassified into class 2 as well. It would seem that class 8, this collection of unarmed anthropomorph figures which do not fit elsewhere, is visually too heterogeneous. It is probable that this causes the poor classification performance for class 21 as well, as this class also contains a variety of otherwise unclassified figures. Figure 9.5 shows examples from class 21. We

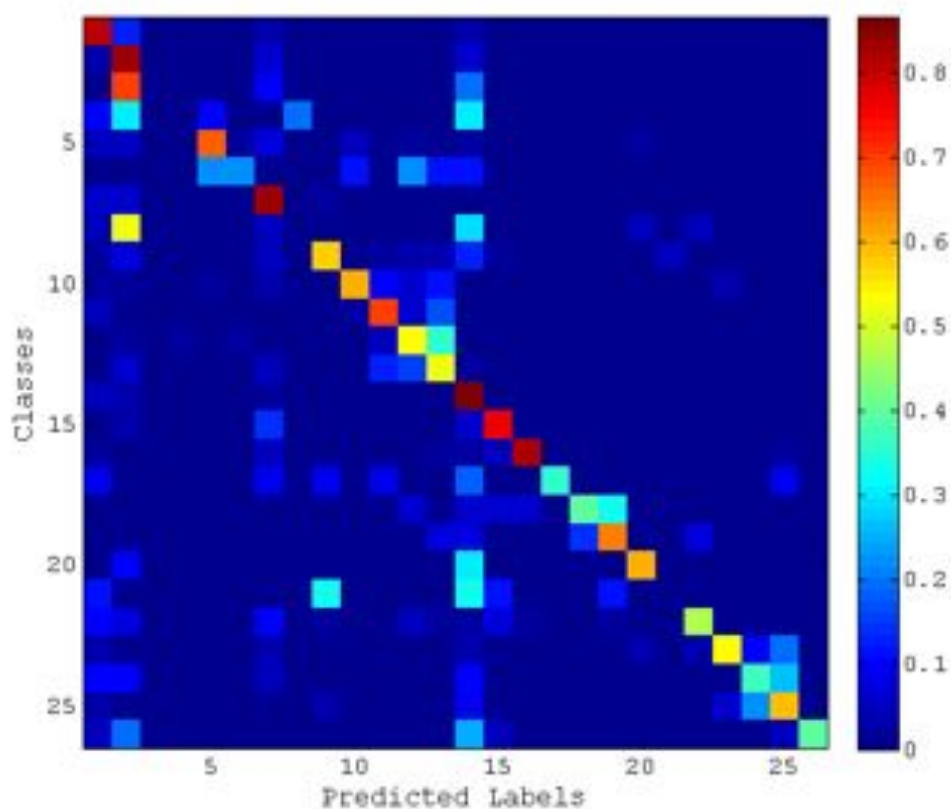


Figure 9.3: Heatmap of the confusion matrix of the classification results of the 26 classes based on the modified Sansoni typology. Each line in the matrix represents the classifications of a certain class. Please refer to Table 2.1 for the class descriptions. See Table 9.1 for percentage values.



Figure 9.4: Examples of modified Sansoni typology class 16 - Axe.

9. AUTOMATED CLASSIFICATION OF PETROGLYPH SHAPES

Table 9.1: Confusion matrix of the classification results of the 26 classes based on the modified Sansoni typology. Each line in the matrix states the percentage of the classifications of a certain class. Please refer to Table 2.1 for the class descriptions. The sum of the values of each line may differ from 100 due to the rounding of every single percentage value.

Classes	Predicted Labels																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1	80	13	0	0	0	0	3	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0
2	3	83	0	0	0	0	6	0	0	0	0	0	0	7	1	0	0	0	0	0	0	0	1	0	0	0
3	0	70	0	0	0	0	10	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0
4	10	30	0	0	10	0	0	20	0	0	0	0	0	30	0	0	0	0	0	0	0	0	0	0	0	0
5	5	5	0	0	67	3	8	0	0	5	0	3	0	3	0	0	0	0	0	3	0	0	0	0	0	0
6	0	0	0	0	22	22	0	0	0	11	0	22	11	11	0	0	0	0	0	0	0	0	0	0	0	0
7	6	6	0	0	0	0	83	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
8	5	52	0	0	0	0	5	0	0	0	0	0	0	29	0	0	0	0	0	5	0	5	0	0	0	0
9	2	7	0	0	0	0	5	0	58	0	2	4	4	12	2	0	0	0	0	0	5	0	0	0	0	0
10	2	2	0	0	2	0	3	0	0	60	10	7	12	0	0	0	0	0	0	0	0	0	3	0	0	0
11	5	0	0	0	0	0	0	0	2	70	7	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	2	0	2	0	2	0	0	0	0	5	54	36	0	0	0	0	0	0	0	0	0	0	0	0	0
13	1	6	0	0	0	0	5	0	0	1	13	16	52	3	0	1	0	0	0	0	0	0	0	1	0	0
14	4	3	0	0	1	0	1	0	1	0	1	1	0	87	0	1	1	0	0	0	0	0	1	0	0	0
15	0	3	0	0	0	0	14	0	0	0	0	0	0	6	77	0	0	0	0	0	0	0	0	0	0	0
16	2	0	0	0	0	0	5	0	0	0	0	0	2	2	5	81	0	0	0	0	0	0	0	0	2	0
17	9	0	0	0	0	0	9	0	9	0	9	0	0	18	0	0	36	0	0	0	0	0	0	0	9	0
18	0	0	0	0	0	0	0	0	0	0	0	7	0	7	7	7	0	40	33	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	7	7	0	0	0	14	64	0	0	7	0	0	0	0
20	0	10	0	0	0	0	0	0	0	0	0	0	0	30	0	0	0	0	0	60	0	0	0	0	0	0
21	11	0	0	0	0	0	0	0	33	0	0	0	0	33	11	0	0	0	11	0	0	0	0	0	0	0
22	10	8	0	0	0	0	10	0	3	0	0	5	3	3	8	3	0	0	3	0	0	46	0	0	0	0
23	3	0	0	0	0	0	3	0	0	0	0	0	0	3	0	0	0	0	0	3	0	3	53	10	20	0
24	11	11	0	0	0	0	5	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	37	26	0
25	3	0	0	0	0	0	0	0	3	0	0	0	0	8	0	0	0	0	0	0	0	0	6	22	58	0
26	5	20	0	0	0	0	0	0	0	0	0	0	0	25	5	0	0	0	0	0	0	0	0	0	5	40



Figure 9.5: Examples of modified Sansoni typology class 21 - Other unclassified.

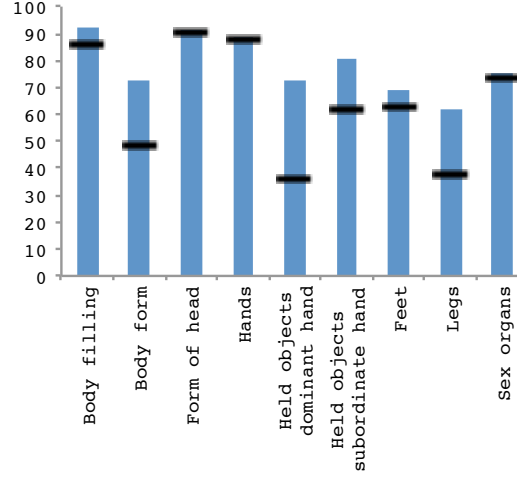


Figure 9.6: Classification accuracy (blue) and baseline (black) for each anthropomorph subdivision of the 3D-Pitoti project typology. Each subdivision reflects a different set of classes (based on the divisions) for all anthropomorph figures. The baseline is the size of the largest class of a subdivision, i.e. the performance of a naive classifier which always predicts the class with the highest *a priori* probability.

observe that the visual similarity is very low.

9.2.1.2 3D-Pitoti Project Typology

As this typology uses subdivisions for the anthropomorph figures (see Subsection 2.6.2), that each classify all anthropomorph figures by different criteria (i.e. any particular anthropomorph will be classified along multiple dimensions in each of the subdivision classes), we first investigate these subdivisions. Figure 9.6 shows the classification performance for each subdivision. We observe that the classification accuracies are relatively high, in a range from 75 to 92%. But, as these subdivisions have only small numbers of classes (from 2 to 7) with sometimes unevenly distributed numbers of examples, we have to consider the baseline classification performance, too. The baseline is the size of the largest class of a subdivision, i.e. the performance of a naive classifier which always predicts the class with the highest *a priori* probability. We observe that for 6 of the 9 subdivisions a performance gain over the baseline can be yielded using our approach. For the remaining 3 subdivisions, our approach yields an accuracy close to the baseline accuracy.

The multiple labeling only applies to anthropomorph figures in the 3D-Pitoti project typology. These are roughly a third of all figures in the dataset. The other two thirds of the figures have a single class label only. Hence, we do not use a fuzzy approach for classification, but intend to apply the classifiers for each subdivision sequentially on

9. AUTOMATED CLASSIFICATION OF PETROGLYPH SHAPES

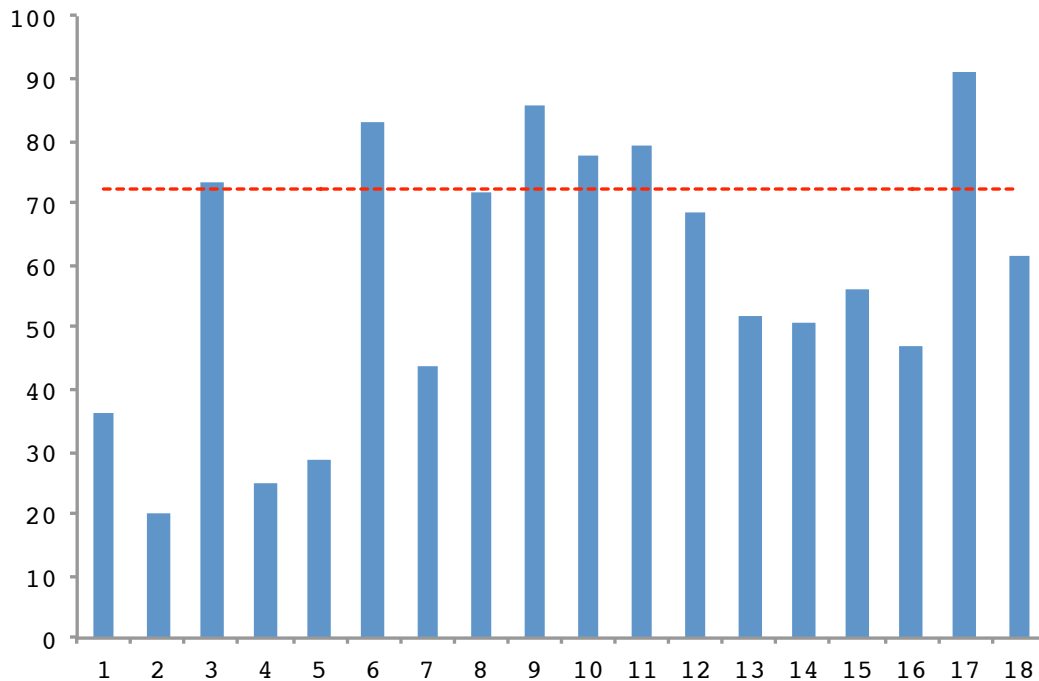


Figure 9.7: Classification accuracy of the classification results of 18 classes based on the 3D-Pitoti project typology. The red line shows the overall accuracy. Please refer to Table 2.2 for the class descriptions of classes 1-16. Classes 17 and 18 are the “body filling” subdivisions of the anthropomorph class.

all anthropomorph candidates resulting from the experiment on the complete 3D-Pitoti project typology. For this classification experiment we include the 2 classes of the best performing subdivision “Body filling”. Hence, we have 18 classes in total. Figure 9.7 shows the classification results. The overall accuracy is 72%, which is clearly better than the 65% of the modified Sansoni typology. We observe that the classes 6 (Cross) and 9 (Shoe/Footprint) yield accuracies over 80%, while class 17 (Anthropomorph with filled body) reaches over 90%. All these classes are highly populated. Class 6 has 118 examples, Class 9 has 104 examples and Class 17 has as many as 284 examples. Column 17 of the confusion matrix in Figure 9.8 shows that many examples of poorer performing classes are falsely classified into class 17. Class 9 attracts a relatively small number of false classifications. Perhaps, then, the examples of this well-performing class have a high visual similarity that suits our classifier well. Figure 9.9 shows examples from class 9. In fact the examples are not visually similar, but they do have comparable contours.

Furthermore, we observe in Figure 9.7 that in contrast to the experiments on the modified Sansoni typology none of the classes yields an accuracy of 0%. Class 2 (Camunian rose, which is the second smallest class) yields the poorest performance - 20%. According to the confusion matrix in Figure 9.8, the misclassifications are mainly into

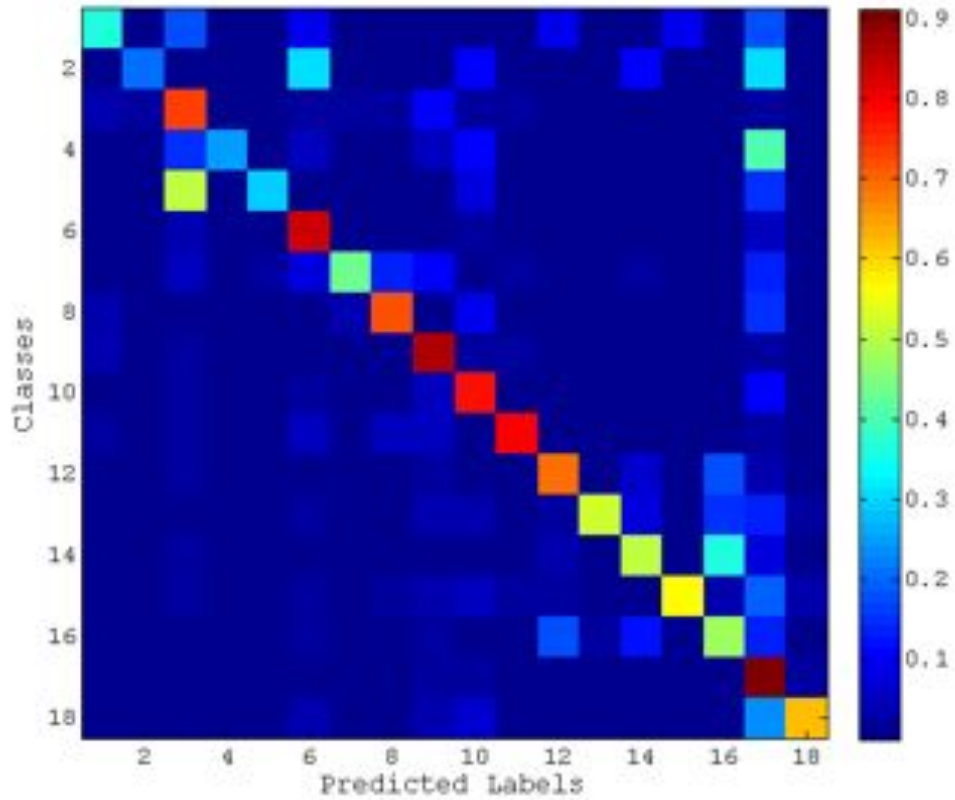


Figure 9.8: Heatmap of the confusion matrix of the classification results of 18 classes based on the 3D-Pitoti project typology. Each line in the matrix represents the classifications of a certain class. Please refer to Table 2.2 for the class descriptions of classes 1-16. Classes 17 and 18 are the “body filling” subdivisions of the anthropomorph class. See Table 9.2 for percentage values.



Figure 9.9: Examples of 3D-Pitoti project typology class 9 - Shoe/Footprint.

9. AUTOMATED CLASSIFICATION OF PETROGLYPH SHAPES

Table 9.2: Confusion matrix of the classification results of 18 classes based on the 3D-Pitoti project typology. Each line in the matrix states the percentage of the classifications of a certain class. Please refer to Table 2.2 for the class descriptions of classes 1-16. Classes 17 and 18 are the “body filling” subdivisions of the anthropomorph class. The sum of the values of each line may differ from 100 due to the rounding of every single percentage value.

Classes	Predicted Labels																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	36	0	18	0	0	9	0	0	0	0	0	9	0	0	9	0	18	0
2	0	20	0	0	0	30	0	0	0	10	0	0	0	10	0	0	30	0
3	4	2	73	0	0	2	2	2	11	2	2	0	0	0	0	0	2	0
4	0	0	15	25	0	5	0	0	5	10	0	0	0	0	0	0	40	0
5	0	0	50	0	29	0	0	0	0	7	0	0	0	0	0	0	14	0
6	1	1	4	0	0	83	0	0	1	3	0	1	0	0	1	0	6	0
7	0	0	5	0	3	8	44	13	10	0	3	0	0	3	0	0	13	0
8	3	0	0	0	0	0	3	71	0	9	0	0	0	0	0	0	14	0
9	3	0	2	0	1	0	0	1	86	2	2	0	0	0	0	1	2	0
10	1	0	3	0	0	2	0	0	5	78	1	0	0	0	0	0	11	0
11	2	0	2	0	0	5	0	5	5	0	79	0	0	0	0	0	2	0
12	0	0	2	0	0	0	0	0	2	0	0	68	0	7	0	18	3	0
13	0	0	0	0	0	2	0	0	3	3	0	2	52	8	0	15	13	2
14	0	0	2	0	0	0	0	0	0	0	0	4	0	51	0	36	7	0
15	0	0	2	0	0	2	0	2	4	5	2	2	0	0	56	4	19	4
16	0	0	1	0	0	2	0	0	2	1	0	18	2	12	0	47	13	1
17	0	0	0	0	0	1	0	0	2	2	0	0	0	0	0	0	91	3
18	0	0	0	0	0	4	0	0	4	6	0	0	0	0	0	0	23	62



Figure 9.10: Examples of 3D-Pitoti project typology class 2 - Camunian rose.

class 6 (cross) and into class 17 (Anthropomorph with filled body), which is the most numerous amongst our classes. Both classes are not visually similar to Class 2, for which Figure 9.10 shows example class members. We assume that the poor performance is due to the fact that the dots around the rose have a negative influence on the pre-processing and skeletonisation of the figure (see Chapter 8).

9.2.2 Details on Single Shape Descriptors and CNNs

In this section we detail the results described in Subsection 9.2.1 with our 3D-Pitoti Project Typology. We investigate the performance of single shape descriptors and several combinations thereof. We add an additional region-based shape descriptor (GFD) and use a state-of-the-art convolutional neural network implementation, in a first step for feature extraction based on an existing network (CNN_IN) and in a second step for re-training of the network with our petroglyph shape data (CNN_RT).

Figure 9.11 shows the results for the evaluation with single descriptors. We observe that the best performing descriptor SC on the small dataset in Subsection 7.3.2 is outperformed by CNN_RT which requires a time-consuming training of the convolutional neural network. The computationally feasible use of a pre-trained network for feature extraction (CNN_IN) yields clearly weaker results than SC. IDSC is behind but is still in the top group of four strong descriptors. The region-based GFD and GHT as well as our skeletal-based GE perform considerably worse than the top group. The performance of the GE is in slight contrast to the performance achieved in Subsection 7.3.2 on the small dataset. We assume that this is a result of the higher number of classes and the more diverse shapes used now.

Figure 9.12 shows the performance of descriptors intermediately fused (see Subsection 7.2.4). We assume, that the results are well comparable with the results in the previous Subsection 9.2.1, where the LOOCV performance of the combination SC, IDSC, GHT, GE on the used 3D-Pitoti -typology is at 72%, whereas the 5-fold cross-validated performance at $k=5$ here yields slightly more 73%. We observe that this descriptor combination which is performing best on the small dataset in Subsection 7.3.2 is clearly outperformed by the combinations that use CNN-based features. Furthermore, we observe that the combination of descriptors clearly improves the performance from 74% with a single descriptor to over 80%.

9.3 Use Scenario

In the previous section we demonstrated that the fully automated classification of petroglyph shapes is feasible and generalises to our larger dataset. Conclusions will be drawn in Section 9.4. In this section we want to look at the results of our approach from a different perspective. A shape classification accuracy of 100% is unlikely to be achieved with real-world data like our petroglyph dataset. Our results (65% and 72% without the use of the re-trained CNN and over 80% with it) cannot be used directly as final classification results but, rather, need manual inspection and, possibly, correction.

9. AUTOMATED CLASSIFICATION OF PETROGLYPH SHAPES

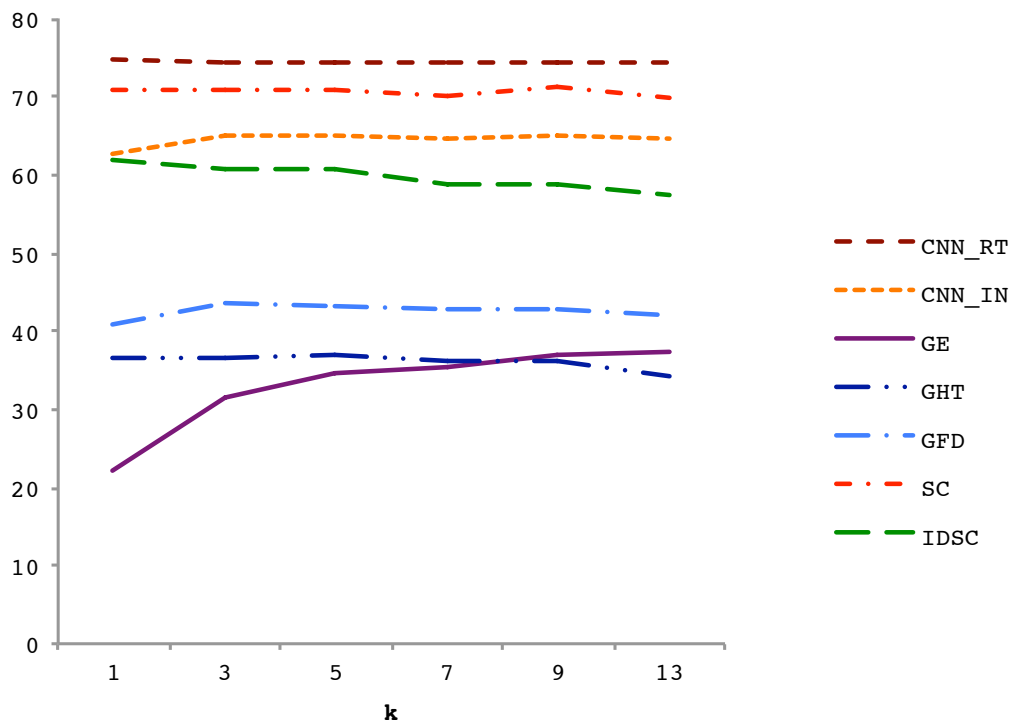


Figure 9.11: Performance of single descriptors in 5-fold cross-validated % of accuracy over number of considered neighbors k . We observe that the best performing descriptor SC (red / long line-dot) on the small dataset in Subsection 7.3.2 is outperformed by CNN_RT (dark red / long line-short line).

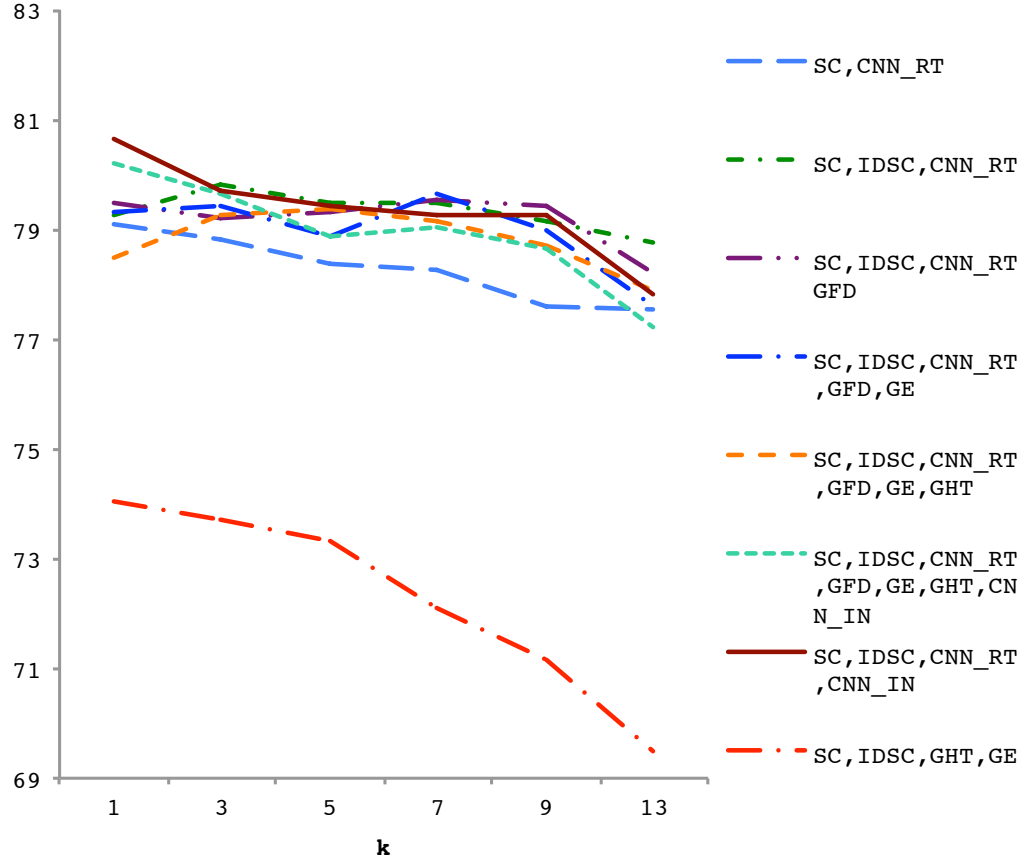


Figure 9.12: Performance of descriptors intermediately fused (see Subsection 7.2.4) in 5-fold cross-validated % of accuracy over number of considered neighbors k . We observe that the combination SC, IDSC, GHT, GE (red / long line-dot) which is performing best on the small dataset in Subsection 7.3.2 and evaluated on two typologies in the previous Section 9.2.1 is clearly outperformed by the combinations that use CNN-based features.

9. AUTOMATED CLASSIFICATION OF PETROGLYPH SHAPES

The question we want to answer now is whether these approaches and results can be applied in a way that usefully supports the work of rock-art researchers. The annotation tool described in Subsection 2.6.3 is a good starting point. Initially this tool was developed to create a dataset and to provide a gold standard for the investigation of petroglyph similarity. However, during the annotation process the tool was found to be of independent value for the archiving, classification and analysis of tracings. The next question is whether - and, if so, how - the automatic classification with its current performance could be included into the manual annotation workflow. For example, we could automatically fill in the predicted class in the drop-downs used for classifying a figure (see Figures 2.21 and 2.25). We could expect correctly pre-selected drop-downs in 65% (modified Sansoni typology), 72% (3D-Pitoti project typology) and 80% (3D-Pitoti project typology with re-trained CNN) of the cases. We could extend this if we not only present a single prediction (the prediction with the highest rank), but if we also present the predictions with lower ranks. Figure 9.13 shows how probable it is that the correct class of a given figure is within the first n items in a ranked list of predicted class candidates. We observe that obviously the values increase the longer the list is. As we want to present the list in an appropriate form in the user interface of the annotation tool, it is advisable to limit it to 5 or 7 items at most. For the 3D-Pitoti project typology, a list size of 5 items yields a performance of 92% (and even 97% with the re-trained CNN) for this semi-automatic classification. For the modified Sansoni typology, the performance is 89%.

From the user perspective this means that around 90% of all classifications of figures can be achieved with a single click, as in all these cases our approach shows the correct class within the five items long ranked list. Furthermore, the sorting of the list also helps the user in making the selection as the highest ranked class candidate is at the top, and more than 80% of the correct classes are within the first three items on the list (see Figure 9.13).

9.4 Conclusion

In this chapter we tested an approach for automated classification of petroglyph shapes on a large real world dataset. Employing a custom-made tool the dataset has been fully expert-annotated according two different typologies containing 18 (3D-Pitoti project) and 26 (modified Sansoni) classes. We have demonstrated, that our previously investigated approach (see Chapter 7) generalises well to a large real world dataset. The overall cross-validated classification accuracies for the two typologies are 72% (even 80% with the re-trained CNN) and 65%, respectively. One general reason for misclassifications is probably the large variation in class sizes. For the 3D-Pitoti project typology, the class sizes vary from 10 to 335, for the modified Sansoni typology from 10 to 143. Weak performing classes are often small classes. Other reasons for weak classification results would appear to be non-homogeneous classes in terms of their members' visual similarity and, perhaps, as-yet-undiscovered issues in our pre-processing pipeline. We have described applications of our annotation tool and demonstrated that the results conse-

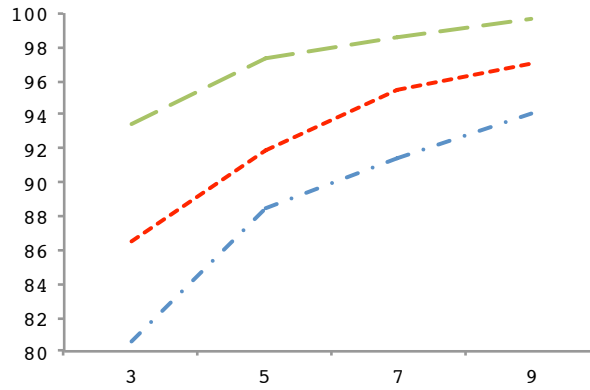


Figure 9.13: Probabilities (y-axis), that the correct class is within the n (x-axis) first predictions of a ranked list of predicted classes. The (red / short line) line shows results for the 3D-Pitoti project typology and the (blue / line-dot) line shows results for the modified Sansoni typology with the descriptor combination performing best on the small dataset in Subsection 7.3.2 and evaluated on two typologies in Subsection 9.2.1. The (green / long line) shows the results for the 3D-Pitoti project typology with the descriptor combination including a CNN classifier (SC, IDSC, CNN_RT; CNN_IN) evaluated in Subsection 9.2.2.

quent upon the application of our classification algorithms to the data developed from said annotation tool can clearly enhance and make more productive the classification work of a rock art researcher.

9. AUTOMATED CLASSIFICATION OF PETROGLYPH SHAPES

Conclusion and Future Work



Figure 10.1: A petroglyph depicting a human figure in Valcamonica, Italy.

IN this thesis, we have developed and described fundamental methods with the aim to enable a digital toolchain for rock art research based on 3D scans. Figure 10.1 shows an example of the artwork. In this chapter we summarise the main technical findings, discuss the potential use of our findings in rock art research and suggest future work.

10. CONCLUSION AND FUTURE WORK

10.1 Conclusion

In the Introduction we defined three ultimate goals for the computational analysis of petroglyphs (see page 2). With the vision of large archives full of rock art images and 3D scans we aim at robust automated methods to

- (a) determine the exact shapes and spatial positions of petroglyphs in images or 3D scans of full rock panels;
- (b) classify the petroglyphs regarding their shapes and pecking styles and
- (c) retrieve similar petroglyphs from different archives of petroglyph images and 3D scans.

In this thesis, we addressed goal (a) and partly goal (b). We developed methods to segment images and 3D scans of rock panels containing petroglyphs. We approached the classification of petroglyphs regarding their shape. Lacking ground truth for pecking style classes, we couldn't work on the classification of pecking styles. We only had access to material from one specific rock art site, thus we couldn't approach goal (c). The chapters of the thesis follow the classic workflow of petroglyph documentation. In Chapters 4 and 5, we presented methods to segment the rock surface in pecked areas and natural rock surface. In Chapter 6, we investigated the pecked regions aiming at automatic distinction of different pecking styles based on their similarity. In Chapters 7 and 8, we approached the numeric description of the shape of petroglyphs based on region, boundary and skeleton to allow similarity estimation of two petroglyphs. In Chapter 9 we showed that the similarity estimation of two petroglyphs allows automated classification in pre-defined classes. The most important findings of our research are:

- Segmentation of rock surface images in pecked areas and natural rock surfaces is feasible based on digital photos. Complementary information of the best performing different texture features is preserved optimally with late fusion.
- Segmentation of rock surfaces based on 3D point clouds outperforms the segmentation based on digital photos. We assumed this result, as information about the micro-topography of the surface which is captured in 3D allows illumination-independent texture description and classification.
- For the segmentation task, 3D descriptors based directly on the surfels (3D points and normals) are outperformed by the employment of 2D visual features on depth maps of the surface which is surprising at first glance. However, we observe that the transfer of the surface's micro-topography to a depth map is possible without any information loss, given sufficient resolution of the depth map and the fact that our surfaces do not contain any self-occlusions. Additionally, the image-space approach can build on a long line of work on 2D visual features.
- 3D descriptors can model different pecking styles well in many cases but fail in others. This task was especially demanding as there is no ground truth for different pecking styles.

- The classification of petroglyphs with respect to their shape is feasible. We have shown, that the combination of region-based, boundary-based and skeletal descriptors yields the best results.
- The features learned by a convolutional neural network (CNN) trained with our data outperform all other shape descriptors we used. But, we show that they contain complementary information which allows to increase classification accuracy when combining the traditional shape descriptors with the CNN features.

For rock art researchers our findings can be useful in several ways. First of all, classic contact tracing to document rock art is a time-consuming task. Additionally the documentation of the pecking style by drawing a point for every single peck mark is highly dependent on the documenting person. Our segmentation method based on 3D scans produces a documentation good enough to require only little corrections by the rock art researcher. Clearly different pecking styles can be separated by the method we propose and together with the 3D material, the rock art researcher can spend more time for interpretation than for documentation. The classification work can clearly be improved as our approach allows a one-click classification for more than 99% of the cases. Given the high number of petroglyphs (more than 300.000 alone in Valcamonica) the automated classification makes the work of rock art researchers more productive. Simple tools can be of use, too. It turned out, that the annotation tool we developed in order to acquire ground truth information for shape classification is still used by the involved rock art researchers even almost two years after the end of the ground truth campaign. They see the value of the tool for organising the large number of classic tracings they work with. We conclude that our methods can improve the work of rock art researchers.

10.2 Future Work

We line out future work in the rock art domain and beyond.

Future work in the rock art domain: We have shown that for petroglyphs 3D surface segmentation clearly outperforms 2D segmentation. 3D surface segmentation requires high-resolution 3D scans. Currently, scans in the required quality cannot be acquired with off-the-shelf hardware but require expensive special hardware and in many cases also post-processing of the captured data. In contrast to that, 2D images can be taken with standard digital cameras. Hence, in cases where 3D scans are not available, 2D segmentation can be an important asset to substitute or at least support the time-consuming tracing. Thus, we see future work in the extension of our 2D approach described in Chapter 4. A larger dataset and the employment of deep learning could improve results.

Our investigation of pecking styles was limited by the fact, that no ground truth for pecking styles exists. Hence, we focused on unsupervised methods. Future work could use a semi-supervised approach, where the user provides hints to a machine learning system. In this way expert knowledge can be integrated and a suitable set of parameters

10. CONCLUSION AND FUTURE WORK

for a given pecking style can be selected in a supervised fashion. This could facilitate the generation of computational models for different pecking styles. The usage of automatic feature selection during model learning (e.g. by sparse SVM or Random Forests) would enable to take the complete set of investigated parameters into account which in turn could maximize the number of pecking style attributes that can be considered in the analysis.

Future work beyond the rock art domain: Our results and findings are relevant beyond the scope of rock art. The most important point deals with the micro-topography of surfaces. We assume that the traditional classification of 2D images of surfaces based on the visual appearance (color and texture) will be superseded by surface classification based on a combination of the tactile appearance and color. The tactile appearance can only be captured with highly detailed 3D scans. Our methods for the distinctive numeric description of the micro-topography of surfaces can be adapted for any kind of surface. As existing 3D descriptors are outperformed by traditional 2D visual features employed on depth maps of the surfaces, future work is necessary for descriptions based directly on point cloud data exploiting the surfels. To decrease computational demand, the surface description can be performed only around salient points. For this, general purpose definitions and detection methods for the saliency of 3D points beyond the peaks and valleys we use are required.

Our investigations on shape similarity have demonstrated that a combination of descriptors that consider boundary, region or skeleton yields very good results. This is of general interest for shape recognition and can - together with our real-world shape dataset - foster new research on shape description. Future work could be done on the evaluation of our descriptor combinations on different shape datasets.

We have demonstrated that our methods can improve the work of rock art researchers and we have shown in our evaluations that we clearly advanced the state-of-the-art. We have argued that our methods are of use beyond rock art in different application domains. We have suggested future work in the rock art domain and beyond. Given the rising availability of 3D scans we expect (and hope!) that our findings will be useful for different applications and will contribute to further methodological development.

List of Figures

1.1	A petroglyph depicting a warrior in Valcamonica, Italy.	1
1.2	Manual contact tracing in Valcamonica, Italy. ©Alberto Marretta, used with permission.	3
1.3	blabla	4
1.4	Simplified system workflow of the complete 3D-Pitoti pipeline. The scanning was carried out as multi-scale approach based on the 3D-reconstruction method Structure-from-Motion (SFM). The massive amount of multi-scale 3D-data required Level-of-Detail (LOD) algorithms for interactive visualisation. The contributions of this thesis (see Section 1.1) are subsumable under “Intelligent Processing”. We thank Alex Kulik for the preparation of the first version of this figure.	5
2.1	A petroglyph scene in Valcamonica, Italy.	9
2.2	Cave painting in the Lascaux cave in France.	11
2.3	A petroglyph depicting a Bubalus in the Sahara in southern Algeria.	12
2.4	Intaglio showing a colibri in Nazca, Peru.	13
2.5	Result of a rubbing documentation. The rubbing shows a rock panel in Tanum.	14
2.6	The process of rubbing on a rock panel in Tanum.	15
2.7	Tracing on a rock in Valcamonica in Italy.	16
2.8	Tracing on a rock in Valcamonica in Italy.	18
2.9	High quality 2D images as used in 2D segmentation. Shadows of trees surrounding the rock are visible.	21
2.10	A high quality 2D image as used in 2D segmentation.	22
2.11	A 3D input point cloud (a) and a close up of the head of the Pitoto (b). The point cloud has 4.4M points. The close up clearly shows the surface morphology of the pecked regions (e.g. in the head and torso of the Pitoto).	23
2.12	An input point cloud (a) with its orthophoto (b) and depth map (c). Brighter values in the depth map indicate higher depth values.	25
2.13	Segmentation ground truth annotation on the 2D image in Photoshop.	26

LIST OF FIGURES

2.14	Orthophotos (a and c) and their corresponding ground truth images (b and d). White areas mark pecked regions in the orthophoto and black areas unpecked regions. The red area surrounding the scan is the area where no information is available.	27
2.15	Petroglyph shapes from our database. The underlying tracings are ©CCSP and Alberto Marretta, used with permission.	28
2.16	Overview of the ground truth annotation web tool interface.	32
2.17	View of a figure in its rock panel tracing.	32
2.18	Use of the polygon selection tool.	33
2.19	Use of the magic wand selection tool.	33
2.20	Comparison of selection without and with filiform selection mode.	34
2.21	Classification process.	35
2.22	Special cases when selecting a figure.	36
2.23	Statistics of numbers of figures in each class. Please note, that the screenshot was taken at a later date than our experiments. As annotation is an ongoing process, the numbers are higher.	37
2.24	All figures for class <i>key</i>	38
2.25	Details for one <i>key</i> figure.	39
3.1	A petroglyph depicting a deer in Valcamonica, Italy.	41
3.2	blabla	43
3.3	bla	44
3.4	Scales of acquisition for the holistic approach of capturing a complete valley in the 3D-Pitoti project.	45
4.1	A human figure and its automatic segmentation result.	49
4.2	Detail of rock art material.	50
4.3	Schema of different fusion strategies applied.	52
4.4	Small details of our test images (see Figures 2.9(a) and 2.10). Problems of the material include grass (tl), shadows (tr), scratches (bl) and cracks (br).	54
4.5	<i>F1</i> - results of single feature classifiers.	56
4.6	<i>F1</i> - results of feature combination with early fusion. The red line marks the result of the best single feature. The corresponding feature combinations are denoted in Table 4.2.	57
4.7	<i>F1</i> - results of feature combination with late fusion. The red line marks the result of the best single feature. The corresponding feature combinations are denoted in Table 4.3.	58
4.8	Two details of our test material with a visualization of late fusion run Y with thresholded majority voting. The grey value of a pixel corresponds to the classification confidence of a pixel. Darker means more confident. We observe that some blurring happens, e.g. at the boundaries of the shield in the upper right. We assume that the blurring is due to the size of the local window we use for feature extraction.	61

LIST OF FIGURES

4.9	Precision and recall for late fusion with thresholded majority voting. The threshold ranges from two to the number of fused classifiers minus one. The corresponding feature combinations are denoted in Table 4.3.	62
4.10	Visualization of two details of our test material classified with different thresholds used in late fusion run Y. Red dots are classified as background. Each column shows the result of a specific threshold t . The most left column corresponds to $t=12$, i.e. all fused classifiers must agree. The most right column corresponds to $t=1$, i.e. at least one classifier must classify a pixel as foreground pixel. We observe that the ideal threshold for segmentation varies.	63
4.11	Our test material overlayed with the best classification result (late fusion run X).	65
5.1	A petroglyph scene with overlayed segmentation ground truth.	67
5.2	blabla	70
5.3	Qualitative evaluation of the discriminative abilities of the employed 3D point cloud geometry-based descriptors. The rows from top to bottom show the ground truth, PFH, FPFH, SHOT and 3DSC. The right hand side shows details. We observe, that each point cloud descriptor produces a large amount of false positives compared to the ground truth in row 1.	72
5.4	Segmentation results with different local point cloud surface descriptors. The left side shows the performance of single features, the right side shows the performance of descriptor combinations. We observe, that the combinations yield only slight improvements.	73
5.5	A 3D surface reconstruction of a rock surface in which the shape of a human figure has been engraved. (a) The colored point cloud; (b) The projected depth map; (c) The deviation map; (d) The enhanced deviation map accentuates well the different surface topographies of the engraving and the surrounding rock surface. Please refer to [ZS15] for details on the process from depth image to enhanced deviation map.	76
5.6	Results of the systematic evaluation with RUSBoost on the complete dataset with 26 3D scans. The enhanced deviation map yields again a significant improvement of performance.	82
6.1	A part of a petroglyph depicting a human with a superimposed arrow containing different pecking styles.	85
6.2	A warrior figure containing two visually clearly distinguishable pecking styles. We observe that the large arrow from left to right has been pecked over the warrior with clearly larger peck marks. This figure serves as a good basic test case whether 3D surface descriptors can help to automatically distinguish between pecking styles.	87
6.3	A scene containing at least two clearly distinguishable pecking styles.	88

LIST OF FIGURES

6.4	Detail of a point cloud with points at peak positions marked red. Please note, that we refer to <i>peaks</i> with respect to the detection method. The peaks are actually valleys if we look at the 3D surface as the detector finds peaks at local maxima in the depth map.	90
6.5	Spherical descriptor footprint with a radius of approx. 2mm. We observe that the peck mark is completely within the footprint.	90
6.6	A petroglyph with homogeneous pecking style.	93
6.7	Visualization of clustering results of densely extracted FPFH descriptors with $r=5\text{mm}$. The figure in the upper left is the original point cloud, the other three figures are the results with 2, 3 and 4 clusters respectively. We observe that the clustering with two clusters is modeling the two visually observed pecking styles well. The clusterings with 3 and 4 clusters do not show improvements.	94
6.8	Visualization of clustering results of densely extracted FPFH descriptors with $r=2\text{mm}$. The figure in the upper left is the original point cloud, the upper right figure shows the result of clustering with two clusters while the figure at the lower left shows the result with three clusters. The two details at the lower right of the figure show the central part of the warrior in the scene. We observe that the superimposition on the figure is clearly distinguishable in the two and three cluster case.	96
6.9	Visualization of clustering results of densely extracted FPFH descriptors with $r=2\text{mm}$. We show results with 2,3,4 and 5 clusters. We observe that the clustering results in vertical stripes that are assumably not caused by the surface topology of the homogeneous pecking style of the petroglyph (see Figure 6.6).	97
6.10	Comparison of different descriptors and footprints. The first row shows clustering results based on the FPFH descriptor with $r=1,2$ and 5mm (from left to right). The second row contains clustering with PFH with $r=1$ and 2mm (from left to right). We observe, that this scan requires a large descriptor footprint as the best result is FPFH with $r=5\text{mm}$. The original scan is shown in Figure 6.2.	98
6.11	A surface detail of a point cloud the points of which have been clustered into two clusters based on PFH with $r=2\text{mm}$. We observe that the flat regions in the surface are in one cluster while the deep peck marks are in the other cluster.	98
6.12	Warrior with surface spheres around peaks serving as basis for topological feature extraction. We observe, that the different pecking styles are modeled well (See Figure 6.2). The third cluster (green) in b) is very small. Please note, that the circular region colored around a peak has only a quarter of the radius that is used as query point neighborhood for feature calculation. We reduced the radius for visual clarity. Overlapping areas of differently colored circles show a mixture of the colors, e.g. yellow in the case of red and green in Fig b).	99

LIST OF FIGURES

6.13	A complex scene with surface spheres around peaks serving as basis for topological feature extraction. We show results with 2 and 3 clusters. We observe, that the different pecking styles are modeled only partly (See Figure 6.3). Please note, that the circular region colored around a peak has only a quarter of the radius that is used as query point neighborhood for feature calculation. We reduced the radius for visual clarity. Overlapping areas of differently colored circles show a mixture of the colors, e.g. yellow in the case of red and green in the left figure.	100
7.1	Petroglyphs with the corresponding smoothed figures, skeletons and the derived graphs.	103
7.2	This figure shows a part of a tracing of a rock in Valcamonica (Coren di Redondo, Rock 1). ©Alberto Marretta, used with permission.	106
7.3	This figure shows the petroglyph dataset that we investigate in this chapter. Each column contains examples of one class. We observe, that some of the classes have high intra-class variance of shape. The petroglyphs are from various rocks in Valcamonica. ©CCSP - Centro Camuno di Studi Preistorici and Alberto Marretta, used with permission.	109
7.4	This figure shows petroglyph tracings (column 1), pre-processed shapes (column 2), extracted skeletons (column 3) and the derived graphs with a pruning threshold of 2, 10 and 30px (columns 4-6). We observe, that the employed skeletonisation algorithm fails to extract details in some cases (heads in rows 1 and 4), while in other cases small details are covered well. Furthermore, we observe, that a high pruning threshold removes skeletal noise (rows 1, 2, 4, 6, 9 and 10), but may also discard relevant topological information (row 5).	110
7.5	This figure shows a sample which is misclassified with GED and GHT. The query image is on the left, and the five nearest neighbors on the right. The first two rows show the result for GED and the utilized graphs. The third row shows the result for GHT.	116
7.6	This figure shows a sample which is correctly classified with GE and misclassified with GHT. The query image is on the left, and the five nearest neighbors on the right. The first two rows show the result for GE and the utilized graphs. The third row shows the result for GHT.	117
8.1	A human figure and a house, the pre-processed shapes and skeletonisation results.	119
8.2	Comparison of the original petroglyph tracing (a) and its skeleton (b). The petroglyph shape pre-processed with the proposed method results in a significantly simplified skeleton that still preserves the important parts of the original shape (c).	120
8.3	Some shapes from example classes (e.g. antropomorph, bird, cross, deer, etc.) from our entire dataset of more than 1100 shapes classified into two different archeological typologies.	124

LIST OF FIGURES

8.4	The workflow of the automated shape pre-processing method. We input a raw binary shape and apply a median filter and binary morphology. . .	127
8.5	Sequence of the automated pre-processing of a shape with the proposed method. The numbers of foreground and background blobs decrease rapidly. At iteration 2, a stable minimum of number of blobs is reached and the stopping criterion aborts the loop at iteration 3. We observe that further iterations decompose the shape starting at iteration 6. . . .	129
8.6	Comparison of skeletons of the original DCE and SPT algorithms (left column) with our improved versions (right column). Our improvements eliminate all shortcomings.	130
8.7	Comparison of histograms of the number of endpoints of the thinning skeletons of unprocessed shapes (left) with the pre-processed shapes (right). The number of endpoints significantly decreases and is mapped to a reasonable range.	133
8.8	Comparison of the results of satisfactory (a) and weak (b) pre-processing of petroglyph shapes.	134
8.9	Examples of successful skeletonisation with all selected and improved algorithms based on shape pre-processing with the proposed method. Note that for DCE and SPT a few spurious branches remain in contrast to thinning and BPR.	135
8.10	Examples of partly erroneous skeletons obtained with the selected algorithms. In all depicted examples spurious branches exist (yellow circles) and for some shapes important parts are deleted (2^{nd} and 3^{rd} row, red circles).	136
9.1	A query petroglyph (left) and the most similar petroglyphs determined by our system.	139
9.2	Classification accuracy of the classification results of the 26 classes based on the modified Sansoni typology. The red line shows the overall accuracy. Please refer to Table 2.1 for the class descriptions.	142
9.3	Heatmap of the confusion matrix of the classification results of the 26 classes based on the modified Sansoni typology. Each line in the matrix represents the classifications of a certain class. Please refer to Table 2.1 for the class descriptions. See Table 9.1 for percentage values.	143
9.4	Examples of modified Sansoni typology class 16 - Axe.	143
9.5	Examples of modified Sansoni typology class 21 - Other unclassified. . .	144
9.6	Classification accuracy (blue) and baseline (black) for each anthropomorph subdivision of the 3D-Pitoti project typology. Each subdivision reflects a different set of classes (based on the divisions) for all anthropomorph figures. The baseline is the size of the largest class of a subdivision, i.e. the performance of a naive classifier which always predicts the class with the highest <i>a priori</i> probability.	145

LIST OF FIGURES

9.7	Classification accuracy of the classification results of 18 classes based on the 3D-Pitoti project typology. The red line shows the overall accuracy. Please refer to Table 2.2 for the class descriptions of classes 1-16. Classes 17 and 18 are the “body filling” subdivisions of the anthropomorph class.	146
9.8	Heatmap of the confusion matrix of the classification results of 18 classes based on the 3D-Pitoti project typology. Each line in the matrix represents the classifications of a certain class. Please refer to Table 2.2 for the class descriptions of classes 1-16. Classes 17 and 18 are the “body filling” subdivisions of the anthropomorph class. See Table 9.2 for percentage values.	147
9.9	Examples of 3D-Pitoti project typology class 9 - Shoe/Footprint.	147
9.10	Examples of 3D-Pitoti project typology class 2 - Camunian rose.	148
9.11	Performance of single descriptors in 5-fold cross-validated % of accuracy over number of considered neighbors k . We observe that the best performing descriptor SC (red / long line-dot) on the small dataset in Subsection 7.3.2 is outperformed by CNN_RT (dark red / long line-short line).	150
9.12	Performance of descriptors intermediately fused (see Subsection 7.2.4) in 5-fold cross-validated % of accuracy over number of considered neighbors k . We observe that the combination SC, IDSC, GHT, GE (red / long line-dot) which is performing best on the small dataset in Subsection 7.3.2 and evaluated on two typologies in the previous Section 9.2.1 is clearly outperformed by the combinations that use CNN-based features.	151
9.13	Probabilities (y-axis), that the correct class is within the n (x-axis) first predictions of a ranked list of predicted classes. The (red / short line) line shows results for the 3D-Pitoti project typology and the (blue / line-dot) line shows results for the modified Sansoni typology with the descriptor combination performing best on the small dataset in Subsection 7.3.2 and evaluated on two typologies in Subsection 9.2.1. The (green / long line) shows the results for the 3D-Pitoti project typology with the descriptor combination including a CNN classifier (SC, IDSC, CNN_RT; CNN_IN) evaluated in Subsection 9.2.2.	153
10.1	A petroglyph depicting a human figure in Valcamonica, Italy.	155
A.1	Orthophoto of scan ID 1.	190
A.2	Orthophoto of scan ID 2.	191
A.3	Orthophoto of scan ID 3.	192
A.4	Orthophoto of scan ID 4.	193
A.5	Orthophoto of scan ID 5.	194
A.6	Orthophoto of scan ID 6.	195
A.7	Orthophoto of scan ID 7.	196
A.8	Orthophoto of scan ID 8.	197
A.9	Orthophoto of scan ID 9.	198
A.10	Orthophoto of scan ID 10.	199

LIST OF FIGURES

A.11 Orthophoto of scan ID 11.	200
A.12 Orthophoto of scan ID 12.	201
A.13 Orthophoto of scan ID 13.	202
A.14 Orthophoto of scan ID 14.	203
A.15 Orthophoto of scan ID 15.	204
A.16 Orthophoto of scan ID 16.	204
A.17 Orthophoto of scan ID 17.	205
A.18 Orthophoto of scan ID 18.	205
A.19 Orthophoto of scan ID 19.	206
A.20 Orthophoto of scan ID 20.	207
A.21 Orthophoto of scan ID 21.	208
A.22 Orthophoto of scan ID 22.	208
A.23 Orthophoto of scan ID 23.	209
A.24 Orthophoto of scan ID 24.	210
A.25 Orthophoto of scan ID 25.	210
A.26 Orthophoto of scan ID 26.	210

List of Tables

2.1	Classes based on branch endpoints of the modified Sansoni typology. . .	30
2.2	Classes based on branch endpoints of the 3D-Pitoti project typology. . .	31
4.1	Features extracted from each local window.	55
4.2	Feature Combinations used for early and late fusion.	57
4.3	Fused classifiers. Please refer to Table 4.2 for features used in feature combinations.	59
5.1	Results of the local 3D descriptors with linear SVM. Numbers provide the median $F1$ as well as the inter quartile range (iqr) in brackets. The PFH descriptor best captures the topographic surface characteristics, however, at the cost of considerable computation time.	78
5.2	Results of the systematic evaluation of all input representations and features with linear SVM on the small dataset (the four point clouds on which the full 3D approach has been evaluated in Section 5.1). Numbers provide the median $F1$ as well as the inter quartile range (iqr) in brackets. ¹ denotes that the result is significantly better than all other results for this feature, i.e. best result of a row. ² denotes that the result is significantly better than all other results for this representation, i.e. best result of a column. GHS+SF applied to the EDM significantly outperform all other features in the experiments.	79
5.3	Results of the systematic evaluation with RUSBoost (same schema as Table 5.2).	81
7.1	Topology based features extracted from the skeletal graphs. Each feature is a description of the whole graph. Please refer to [BdW12] for more details.	112
7.2	Classification accuracy of GED in percent. We use LOOCV-validated 5-NN classification. The maximum number of open paths for the A* beam approximation is 6000.	113
7.3	Classification accuracy (in percent) of GE utilizing single scalar topological features. We use LOOCV-validated 5-NN classification. p denotes the size of the pruning threshold. The three best values are emphasized. Please refer to Table 7.1 for the descriptions corresponding to the feature IDs.	114

LIST OF TABLES

7.4	Classification accuracy (in percent) of GE using combinations of 1 to 10 single topology features that perform best for each dimension. We use LOOCV-validated 5-NN classification with brute-force feature selection. The pruning threshold is 10px. The best values are emphasized. Please refer to Table 7.1 for the descriptions corresponding to the feature IDs. .	115
7.5	Classification accuracy (in percent) of GE, SC, IDSC and GHT and combinations thereof. We use LOOCV-validated 5-NN classification for single shape features and the classifier fusion method described in Section 7.2.4.	115
8.1	Comparison of recent skeletonisation methods with respect to the identified criteria. Note that for point-based pruning approaches the first two criteria do not apply because they do not distinguish between significant and insignificant branches.	125
8.2	Results of the quantitative evaluation of the proposed pre-processing method on the entire dataset.	133
8.3	Evaluation of the four selected skeletonisation algorithms on the entire dataset. BPR and thinning perform best on the pre-processed shapes. .	137
9.1	Confusion matrix of the classification results of the 26 classes based on the modified Sansoni typology. Each line in the matrix states the percentage of the classifications of a certain class. Please refer to Table 2.1 for the class descriptions. The sum of the values of each line may differ from 100 due to the rounding of every single percentage value.	144
9.2	Confusion matrix of the classification results of 18 classes based on the 3D-Pitoti project typology. Each line in the matrix states the percentage of the classifications of a certain class. Please refer to Table 2.2 for the class descriptions of classes 1-16. Classes 17 and 18 are the “body filling” subdivisions of the anthropomorph class. The sum of the values of each line may differ from 100 due to the rounding of every single percentage value.	148
A.1	Summary of the reconstructed surfaces in the dataset. For each surface we provide an ID, description, its location (in terms of site, rock, and area) and the reconstruction method.	188
A.2	Overview of basic measures of the digitized surfaces: the covered area (in pixels at 300dpi and in cm ²), the number of 3D points in the point cloud, the percentage of pecked regions, the number of disconnected pecked regions and the range of depth values.	189

References

- [AB96] Carlo Arcelli and Gabriella Sanniti di Baja. Skeletons of planar patterns. In T. Yung Kong and Azriel Rosenfeld, editors, *Topological Algorithms for Digital Image Processing*, volume 19 of *Machine Intelligence and Pattern Recognition*, pages 99 – 143. North-Holland, 1996.
- [ABR⁺14] M. Aubert, A. Brumm, M. Ramli, T. Sutikna, E. W. Saptomo, B. Hakim, M. J. Morwood, G. D. van den Bergh, L. Kinsley, and A. Dosseto. Pleistocene cave art from Sulawesi, Indonesia. *Nature*, 514(7521):223–227, October 2014.
- [AC94] Emmanuel Anati and Tiziana Cittadini. *Valcamonica rock art: A new history for Europe*. Edizioni del Centro, Capo di Ponte, Italy, 1994.
- [AEET08] Cagri Aslan, Aykut Erdem, Erkut Erdem, and Sibel Tari. Disconnected skeleton: shape at its absolute scale. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(12):2188–2203, 2008.
- [AH95] H. Anys and Dong-Chen He. Evaluation of textural and multipolarization radar features for crop classification. *IEEE Transactions on Geoscience and Remote Sensing*, 33(5):1170 –1181, sep 1995.
- [Ana76] Emmanuel Anati. *Evolution and style in Camunian rock art: An inquiry into the formation of European civilization*. Edizioni del Centro, Capo di Ponte, Italy, 1976.
- [APR15] Craig Alexander, Axel Pinz, and Christian Reinbacher. Multi-scale 3d rock-art recording. *Digital Applications in Archaeology and Cultural Heritage*, 2(2–3):181 – 195, 2015.
- [ASM96] ASME. *ASME B46.1-1995 Surface texture (surface roughness, waviness and lay): An American National Standard*. The American Society of Mechanical Engineers (ASME), New York, USA, 1996.
- [Baj06] Gabriella Sanniti di Baja. Skeletonization of digital objects. In *Proceedings of the 11th Iberoamerican conference on Progress in Pattern*

REFERENCES

- Recognition, Image Analysis and Applications (CIARP06)*, pages 1–13. Springer DE, 2006.
- [BdW12] Gergana Bounova and Olivier de Weck. Overview of metrics and their correlation patterns for multiple-metric topology analysis on heterogeneous graph ensembles. *Phys. Rev. E*, 85:016117, Jan 2012.
- [BET09] Emre Baseski, Aykut Erdem, and Sibel Tari. Dissimilarity between two skeletal trees in a context. *Pattern Recognition*, 42(3):370–385, 2009.
- [BHH78] George E.P. Box, William Gordon Hunter, and J. Stuart Hunter. *Statistics for experimenters*. John Wiley and sons New York, 1978.
- [BJ03] Liam Blunt and Xiang Jiang. *Advanced techniques for assessment surface topography: development of a basis for 3D surface texture standards "surfstand"*. Elsevier, 2003.
- [BL07] Xiang Bai and Longin Jan Latecki. Discrete skeleton evolution. In *Proceedings of the 6th Int. Conf. on Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 362–374, Berlin, Heidelberg, 2007. Springer-Verlag.
- [BL08] Xiang Bai and L.J. Latecki. Path similarity skeleton graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(7):1282–1292, July 2008.
- [BLL07] Xiang Bai, Longin Latecki, and Wen-yu Liu. Skeleton pruning by contour partitioning with discrete curve evolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(3):449–462, March 2007. 00228.
- [BLT09] Xiang Bai, Wenyu Liu, and Zhuowen Tu. Integrating contour and skeleton for shape classification. In *2009 IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops)*, page 360–367. IEEE, 2009.
- [Blu67] H. Blum. A transformation for extracting new descriptors of shape. In *Models for the Perception of Speech and Visual Form, Proceedings of Meeting held in Boston, Nov. 1964*, pages 362–380, Cambridge, 1967. MIT Press.
- [BMdA12] J.A. Barcelo and V. Moitinho de Almeida. Functional analysis from visual and non-visual data. an artificial intelligence approach. *Mediterranean Archaeology and Archaeometry*, 12(2):273–321, 2012.

-
- [BMP02] Serge Belongie, Jitendra Malik, and Jan Puzicha. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(4):509–522, 2002.
- [BN78] H. Blum and Roger N. Nagel. Shape description using weighted symmetric axis features. *Pattern Recognition*, 10(3):167–180, 1978.
- [Cla02] David A Clausi. An analysis of co-occurrence texture statistics as a function of grey level quantization. *Canadian Journal of remote sensing*, 28(1):45–62, 2002.
- [CV95] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.
- [DABRR06] Margarita Díaz-Andreu, Christopher Brooke, Michael Rainsbury, and Nick Rosser. The spiral that vanished: the application of non-contact recording techniques to an elusive rock art motif at Castlerigg stone circle in Cumbria. *Journal of Archaeological Science*, 33(11):1580–1587, November 2006. 00034.
- [dARBP13] V. M. de Almeida, R. Rosillo, J. A. Barcel’o, and A. Palomo. Linking 3d digital surface texture data with ancient manufacturing procedures. In *Digital Heritage International Congress (DigitalHeritage), 2013*, volume 1, pages 735–738, Oct 2013.
- [Din55] G. P. Dinneen. Programming pattern recognition. In *Proceedings of the March 1-3, 1955, Western Joint Comp. Conf.*, AFIPS ’55 (Western), pages 94–100, NY, USA, 1955. ACM.
- [DJLW08] Ritendra Datta, Dhiraj Joshi, Jia Li, and James Z Wang. Image retrieval: Ideas, influences, and trends of the new age. *ACM Comput. Surv.*, 40(2):5:1–5:60, May 2008.
- [DK12] Tal Darom and Yosi Keller. Scale-invariant features for 3-d mesh models. *IEEE Transactions on Image Processing*, 21(5):2758–2769, 2012.
- [DP13] Vincenzo Deufemia and Luca Paolino. Combining unsupervised clustering with a non-linear deformation model for efficient petroglyph recognition. In George Bebis, Richard Boyle, Bahram Parvin, Darko Koracin, Baoxin Li, Fatih Porikli, Victor Zordan, James Klosowski, Sabine Coquillart, Xun Luo, Min Chen, and David Gotz, editors, *Advances in Visual Computing*, number 8034 in Lecture Notes in Computer Science, pages 128–137. Springer Berlin Heidelberg, January 2013.

REFERENCES

- [DP14] Vincenzo Deufemia and Luca Paolino. Segmentation and recognition of petroglyphs using generic fourier descriptors. In Abderrahim Elmoataz, Olivier Lezoray, Fathallah Nouboud, and Driss Mammass, editors, *Image and Signal Processing*, volume 8509 of *Lecture Notes in Computer Science*, pages 487–494. Springer International Publishing, 2014.
- [DPdL12] V. Deufemia, L. Paolino, and H. de Lumley. Petroglyph recognition using self-organizing maps and fuzzy visual language parsing. In *2012 IEEE 24th International Conference on Tools with Artificial Intelligence (ICTAI)*, volume 1, pages 852–859, 2012.
- [DR04] Cecilia Di Ruberto. Recognition of shapes by attributed skeletal graphs. *Pattern Recognition*, 37(1):21–31, 2004.
- [DSS92] WP Dong, PJ Sullivan, and KJ Stout. Comprehensive study of parameters for characterizing three-dimensional surface topography i: Some inherent properties of parameter variation. *Wear*, 159(2):161–171, 1992.
- [DSS93] WP Dong, PJ Sullivan, and KJ Stout. Comprehensive study of parameters for characterizing three-dimensional surface topography ii: Statistical properties of parameter variation. *Wear*, 167(1):9–21, 1993.
- [DSS94a] WP Dong, PJ Sullivan, and KJ Stout. Comprehensive study of parameters for characterising three-dimensional surface topography: Iii: Parameters for characterising amplitude and some functional properties. *Wear*, 178(1):29–43, 1994.
- [DSS94b] WP Dong, PJ Sullivan, and KJ Stout. Comprehensive study of parameters for characterising three-dimensional surface topography: Iv: parameters for characterising spatial and hybrid properties. *Wear*, 178(1):45–60, 1994.
- [DT05] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893. IEEE, 2005.
- [DvGNK99] Kristin J Dana, Bram van Ginneken, Shree K Nayar, and Jan J Koenderink. Reflectance and texture of real-world surfaces. *ACM Trans. Graph.*, 18(1):1–34, January 1999.
- [DvLV08] M. Fatih Demirci, Reinier H. van Leuken, and Remco C. Veltkamp. Indexing through Laplacian spectra. *Computer Vision and Image Understanding*, 110(3):312–325, June 2008.

-
- [ET10] Aykut Erdem and Sibel Tari. A similarity-based approach for shape classification using Aslan skeletons. *Pattern Recognition Letters*, 31(13):2024–2032, 2010.
- [FHK⁺04] Andrea Frome, Daniel Huber, Ravi Kolluri, Thomas Bülow, and Jitendra Malik. Recognizing objects in range data using regional point descriptors. In *Computer Vision-ECCV 2004*, pages 224–237. Springer, 2004. 00401.
- [Fis35] Ronald Aylmer Fisher. *The design of experiments*. Oliver & Boyd, 1935.
- [FS⁺96] Yoav Freund, Robert E Schapire, et al. Experiments with a new boosting algorithm. In *In Proceedings of the International Conference on Machine Learning*, volume 96, pages 148–156, 1996.
- [GAMNRGM11] Diego Gonzalez-Aguilera, Angel Muñoz-Nieto, Pablo Rodriguez-Gonzalvez, and Mario Menéndez. New tools for rock art modelling: automated sensor integration in Pindal Cave. *Journal of Archaeological Science*, 38(1):120–128, January 2011. 00013.
- [GVB12] Jaume Gibert, Ernest Valveny, and Horst Bunke. Graph embedding in vector spaces by node attribute statistics. *Pattern Recogn.*, 45(9):3072–3083, September 2012.
- [HD84] S.B. Ho and C.R. Dyer. Medial-axis based shape smoothing. Technical Report 557, University of Wisconsin-Madison Department of Computer Sciences, 1984.
- [How04] Nicholas R. Howe. Code implementations by Nicholas R. Howe, 2004.
- [HR08] T. Hengl and H.I. Reuter, editors. *Geomorphometry: Concepts, Software, Applications*, volume 33. Elsevier, Amsterdam, 2008.
- [HSD73] Robert M. Haralick, K. Shanmugam, and Its’Hak Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6):610–621, November 1973.
- [II02] ISO-IEC. *Information Technology - Multimedia Content Description Interface*. 15938. ISO/IEC, Moving Pictures Expert Group, 1st edition, 2002.
- [JH99] Andrew E. Johnson and Martial Hebert. Using spin images for efficient object recognition in cluttered 3d scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5):433–449, 1999.

REFERENCES

- [JSD⁺14] Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, and Trevor Darrell. Caffe: Convolutional architecture for fast feature embedding. *arXiv preprint arXiv:1408.5093*, 2014.
- [KAH⁺13] Walter G. Kropatsch, Nicole M. Artner, Yll Haxhimusa, Xiaoyi Jiang, David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Doug Tygar, Moshe Y. Vardi, and Gerhard Weikum, editors. *Graph-Based Representations in Pattern Recognition*, volume 7877 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.
- [KC09] Stelios Krinidis and V. Chatzis. A skeleton family generator via physics-based deformable models. *IEEE Transactions on Image Processing*, 18(1):1–11, 2009.
- [KCRU58] R. A. Kirsch, L. Cahn, C. Ray, and G. H. Urban. Experiments in processing pictorial information with a digital computer. In *Papers and Discussions Presented at the December 9-13, 1957, Eastern Joint Comp. Conf.: Computers with Deadlines to Meet*, IRE-ACM-AIEE '57 (Eastern), pages 221–229, New York, USA, 1958. ACM.
- [Kir79] David Kirkpatrick. Efficient computation of continuous skeletons. In *Proceedings of the 20th Annual Symposium on Foundations of Computer Science, 1979.*, pages 18–27, San Juan, Puerto Rico, 1979. IEEE.
- [KJPK02] Kwang In Kim, Keechul Jung, Se Hyun Park, and Hang Joon Kim. Support vector machines for texture classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:1542–1550, 2002.
- [KK13] Stelios Krinidis and Michail Krinidis. Empirical mode decomposition on skeletonization pruning. *Image and Vision Computing*, 31(8):533 – 541, 2013.
- [KOW08] Hee-Kooi Khoo, Hong-Choon Ong, and Ya-Ping Wong. Image texture classification using combined grey level Co-Occurrence probabilities and support vector machines. In *Proceedings of the 2008 Fifth International Conference on Computer Graphics, Imaging and Visualization*, page 180–184, Washington, DC, USA, 2008. IEEE Computer Society.
- [KS08] Karthik Krish and Wesley Snyder. A new accumulator-based approach to shape recognition. In *Proceedings of the 4th International Symposium on Advances in Visual Computing, Part II*, ISVC '08, page 157–169, Berlin, Heidelberg, 2008. Springer-Verlag.

-
- [KSK01] Philip N. Klein, Thomas B. Sebastian, and Benjamin B. Kimia. Shape matching using edit-distance: An implementation. In *Proceedings of the Twelfth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '01, pages 781–790, Philadelphia, PA, USA, 2001. Society for Industrial and Applied Mathematics.
- [Kun04] Ludmila I. Kuncheva. *Combining Pattern Classifiers: Methods and Algorithms*. Wiley-Interscience, 2004.
- [Kva13] Kenneth Kvamme. An examination of automated archaeological feature recognition in remotely sensed imagery. *Computational Approaches to Archaeological Spaces*, 60:53, 2013.
- [Lea11] Richard Leach. *Optical measurement of surface topography*. Springer, 2011.
- [LJ07] Haibin Ling and David W Jacobs. Shape classification using the inner-distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(2):286–299, 2007.
- [LL99] Longin Jan Latecki and Rolf Lakämper. Polygon evolution by vertex deletion. In *Proceedings of the Second International Conference on Scale-Space Theories in Computer Vision*, pages 398–409. Springer, 1999.
- [LL00] Longin Jan Latecki and Rolf Lakamper. Shape similarity measure based on correspondence of visual parts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1185–1190, 2000.
- [LLE00] Longin Jan Latecki, Rolf Lakamper, and T. Eckhardt. Shape descriptors for non-rigid shapes with a single closed contour. In *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, volume 1, page 424–429. IEEE, 2000.
- [LM99] Thomas Leung and Jitendra Malik. Recognizing surfaces using Three-Dimensional textons. In *Proceedings of the International Conference on Computer Vision-Volume 2*, ICCV '99, page 1010–1025, Washington, DC, USA, 1999. IEEE Computer Society.
- [LNCV10] José Luis Lerma, Santiago Navarro, Miriam Cabrelles, and Valentín Villaverde. Terrestrial laser scanning and close range photogrammetry for 3d archaeological documentation: the Upper Palaeolithic Cave of Parpalló as a case study. *Journal of Archaeological Science*, 37(3):499–507, March 2010. 00090.
- [Lon98] Sven Loncaric. A survey of shape analysis techniques. *Pattern recognition*, 31(8):983–1001, 1998.

REFERENCES

- [Low04] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, November 2004.
- [LPC⁺00] Marc Levoy, Kari Pulli, Brian Curless, Szymon Rusinkiewicz, David Koller, Lucas Pereira, Matt Ginzton, Sean Anderson, James Davis, Jeremy Ginsberg, and others. The digital Michelangelo project: 3d scanning of large statues. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 131–144. ACM Press/Addison-Wesley Publishing Co., 2000. 01775.
- [LS06] George V. Landon and W. Brent. Seales. Petroglyph digitization: enabling cultural heritage scholarship. *Machine Vision and Applications*, 17(6):361–371, October 2006.
- [LSYZ11] Geng Li, Murat Semerci, Bülent Yener, and Mohammed J. Zaki. Graph classification via topological and label attributes. In *9th Workshop on Mining and Learning with Graphs (with SIGKDD)(August 2011)*, 2011. 00009.
- [LWH⁺12] Hongzhi Liu, Zhonghai Wu, D. Frank Hsu, Bradley S. Peterson, and Dongrong Xu. On the generation and pruning of skeletons using generalized voronoi diagrams. *Pattern Recognition Letters*, 33(16):2113–2119, December 2012.
- [LWZH13] Hongzhi Liu, Zhong-Hai Wu, Xing Zhang, and D. Frank Hsu. A skeleton pruning algorithm based on information fusion. *Pattern Recognition Letters*, 34(10):1138 – 1145, 2013.
- [MAK⁺97] Farzin Mokhtarian, Sadegh Abbasi, Josef Kittler, et al. Efficient and robust retrieval by shape content through curvature scale space. *Series on Software Engineering and Knowledge Engineering*, 8:51–58, 1997.
- [Mar11] Alberto Marretta. Valcamonica Rock Art: An Extraordinary Archaeological Source. In Alberto Marretta and Tiziana Cittadini, editors, *Valcamonica rock-art parks: guide to visiting routes*, pages 8–29. Edizione del Centro, Capo di Ponte, 1 edition, 2011.
- [MCH10] Fei Mai, C. Q. Chang, and Y. S. Hung. Affine-invariant shape matching and recognition under partial occlusion. In *17th IEEE International Conference on Image Processing (ICIP)*, page 4605–4608. IEEE, 2010.
- [MM86] Farzin Mokhtarian and Alan Mackworth. Scale-based description and recognition of planar curves and two-dimensional shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):34–43, 1986.

-
- [Mok95] Farzin Mokhtarian. Silhouette-based isolated object recognition through curvature scale space. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(5):539–544, 1995.
- [Mon68] U. Montanari. A method for obtaining skeletons using a quasi-euclidean distance. *Journal of the ACM (JACM)*, 15(4):600–624, October 1968.
- [Mon69] Ugo Montanari. Continuous skeletons from digitized images. *Journal of the ACM (JACM)*, 16(4):534–549, October 1969.
- [OI92] Robert L. Ogniewicz and M. Ilg. Voronoi skeletons: Theory and applications. In *Proceedings of the IEEE Conf. on Comp. Vision and Patt. Rec. (CVPR)*, pages 63–69, 1992.
- [OPH96] Timo Ojala, Matti Pietikäinen, and David Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 29(1):51–59, 1996.
- [OPM02] Timo Ojala, Matti Pietikainen, and Topi Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, 2002.
- [Par11] J. R Parker. *Algorithms for image processing and computer vision*. Wiley Publishing, Inc, Indianapolis, USA, 2nd edition edition, 2011.
- [Pav78] Theodosios Pavlidis. A review of algorithms for shape analysis. *Computer graphics and image processing*, 7(2):243–258, 1978.
- [PCGV02] Mari Partio, Bogdan Cramariuc, Moncef Gabbouj, and Ari Visa. Rock texture retrieval using gray level Co-Occurrence matrix. In *Proceedings of the 5th Nordic Signal Processing Symposium*, Oslo, 2002.
- [PDM02] Euripides G. M. Petrakis, Aristeidis Diplaros, and Evangelos Milios. Matching and retrieval of distorted and occluded shapes using dynamic programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(11):1501–1516, 2002.
- [PHGD⁺12] A. W. G. Pike, D. L. Hoffmann, M. Garcia-Diez, P. B. Pettitt, J. Alcolea, R. De Balbin, C. Gonzalez-Sainz, C. de las Heras, J. A. Lasheras, R. Montes, and J. Zilhao. U-Series Dating of Paleolithic Art in 11 Caves in Spain. *Science*, 336(6087):1409–1413, June 2012.
- [Rao05] C Rao. Linear model selection by cross-validation. *Journal of Statistical Planning and Inference*, 128(1):231–240, 2005.

REFERENCES

- [RBB09] Radu Bogdan Rusu, Nico Blodow, and Michael Beetz. Fast point feature histograms (FPFH) for 3d registration. In *Proceedings of IEEE International Conference on Robotics and Automation, 2009. ICRA'09.*, pages 3212–3217. IEEE, 2009. 00320.
- [RC11] Radu Bogdan Rusu and Steve Cousins. 3D is here: Point Cloud Library (PCL). In *IEEE International Conference on Robotics and Automation (ICRA)*, Shanghai, China, May 9-13 2011.
- [REB13] Kaspar Riesen, Sandro Emmenegger, and Horst Bunke. A novel software toolkit for graph edit distance computation. In *Graph-Based Representations in Pattern Recognition*, pages 142–151. Springer, 2013.
- [RMBB08] Radu Bogdan Rusu, Zoltan Csaba Marton, Nico Blodow, and Michael Beetz. Persistent point feature histograms for 3d point clouds. In *Proc 10th Int Conf Intel Autonomous Syst (IAS-10)*, Baden-Baden, Germany, pages 119–128, 2008. 00035.
- [RRPOGP11] Edgar Roman-Rangel, Carlos Pallan, Jean-Marc Odobez, and Daniel Gatica-Perez. Analyzing ancient maya glyph collections with contextual shape descriptors. *International Journal of Computer Vision*, 94(1):101–117, 2011.
- [SAC07] Mark D Smucker, James Allan, and Ben Carterette. A comparison of statistical significance tests for information retrieval evaluation. In *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*, pages 623–632. ACM, 2007.
- [SB98] Doron Shaked and Alfred M. Bruckstein. Pruning medial axes. *Computer Vision Image Understanding*, 69(2):156–169, February 1998.
- [SB11] Markus Seidl and Christian Breiteneder. Detection and classification of petroglyphs in gigapixel images – preliminary results. *VAST11: The 12th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage - Short Papers*, pages 45–48, 2011.
- [SB12] Markus Seidl and Christian Breiteneder. Automated petroglyph image segmentation with interactive classifier fusion. In *Proceedings of the Eighth Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP '12)*, 2012.
- [SBH⁺11] Wei Shen, Xiang Bai, Rong Hu, Hongyuan Wang, and Longin Latecki. Skeleton growing and pruning with bending potential ratio. *Pattern Recognition*, 44(2):196–209, February 2011.

-
- [SBYL13] Wei Shen, Xiang Bai, XingWei Yang, and Longin Jan Latecki. Skeleton pruning as trade-off between skeleton simplicity and reconstruction error. *Science China Information Sciences*, 56(4):1–14, January 2013.
- [SCC⁺11] Roberto Scopigno, Marco Callieri, Paolo Cignoni, Massimiliano Corsini, Matteo Dellepiane, Federico Ponchio, and Guido Ranzuglia. 3d models for cultural heritage: Beyond plain visualization. *IEEE Computer*, 44(7):48–55, 2011.
- [Sco12] Roberto Scopigno. Sampled 3d models for cultural heritage: which uses beyond visualization? *Virtual Archaeology Review*, 3(5):109–115, 2012.
- [SDR⁺10] Juan Ortiz Sanz, Maria de la Luz Gil Docampo, Santiago Martínez Rodríguez, María Teresa Rego Sanmartín, and Gonzalo Meijide Camelelle. A simple methodology for recording petroglyphs using low-cost digital image correlation photogrammetry and consumer-grade digital cameras. *Journal of Archaeological Science*, 37(12):3158–3169, December 2010. 00026.
- [SG09] U. Sansoni and S. Gavaldo. *Lucus rupestris: sei millenni d’arte rupestre a Campanine di Cimbergo*. Edizioni del Centro, Capo di Ponte (BS), Italy, 2009.
- [Sil09] Roland Silva. *Rock Art Sites on the UNESCO World Heritage List*. ICOMOS, International Council on Monuments and Sites, Paris, 2009.
- [SK96] K. Siddiqi and B.B. Kimia. A shock grammar for recognition. In *Proceedings of 1996 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR ’96*,, pages 507–513, 1996.
- [SKVHN08] Chris Seiffert, Taghi M Khoshgoftaar, Jason Van Hulse, and Amri Napolitano. Rusboost: Improving classification performance when training data is skewed. In *19th International Conference on Pattern Recognition, 2008. ICPR 2008*, pages 1–4. IEEE, 2008.
- [SKVHN10] Chris Seiffert, Taghi M Khoshgoftaar, Jason Van Hulse, and Amri Napolitano. Rusboost: A hybrid approach to alleviating class imbalance. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 40(1):185–197, 2010.
- [SRKB11] Bastian Steder, Radu Bogdan Rusu, Kurt Konolige, and Wolfram Burgard. Point feature extraction on 3d range scans taking into account object boundaries. In *Proceedings of 2011 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2601–2608. IEEE, 2011.

REFERENCES

- [SS05] Kang B Sun and Boaz J Super. Classification of contour shapes using class segment sets. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, page 727–733. IEEE, 2005.
- [SSDZ99] Kaleem Siddiqi, Ali Shokoufandeh, SvenJ. Dickinson, and StevenW. Zucker. Shock graphs and shape matching. *International Journal of Computer Vision*, 35(1):13–32, 1999.
- [ST99] L-K Soh and Costas Tsatsoulis. Texture analysis of sar sea ice imagery using gray level co-occurrence matrices. *IEEE Transactions on Geoscience and Remote Sensing*, 37(2):780–795, 1999.
- [STDS14] Samuele Salti, Federico Tombari, and Luigi Di Stefano. SHOT: Unique signatures of histograms for surface and texture description. *Computer Vision and Image Understanding*, 125:251–264, August 2014.
- [SWA15] Markus Seidl, Ewald Wieser, and Craig Alexander. Automated classification of petroglyphs. *Digital Applications in Archaeology and Cultural Heritage*, 2(2,3):196 – 212, 2015.
- [SWZ⁺15] Markus Seidl, Ewald Wieser, Matthias Zeppelzauer, Axel Pinz, and Christian Breiteneder. Graph-Based Shape Similarity of Petroglyphs. In Lourdes Agapito, Michael M. Bronstein, and Carsten Rother, editors, *Computer Vision - ECCV 2014 Workshops*, volume 8925 of *Lecture Notes in Computer Science*, pages 133–148. Springer International Publishing, 2015.
- [Ted00] Laura Anne Tedesco. Lascaux (ca. 15,000 B.C.). In *Heilbrunn Timeline of Art History*. The Metropolitan Museum of Art, New York, October 2000.
- [Tel12] Alexandru Telea. Feature preserving smoothing of shapes using saliency skeletons. In *Vis. in Medicine and Life Sciences II*, pages 153–170. Springer, 2012.
- [TJ98] Mihran Tuceryan and Anil K Jain. Texture analysis. *The handbook of pattern recognition and computer vision*, 2:207–248, 1998.
- [TLS09] Øivind Due Trier, Siri Øyen Larsen, and Rune Solberg. Automatic detection of circular structures in high-resolution satellite images of agricultural land. *Archaeological Prospection*, 16(1):1–15, 2009.
- [TSDS10] Federico Tombari, Samuele Salti, and Luigi Di Stefano. Unique signatures of histograms for local surface description. In *Computer Vision-ECCV 2010*, pages 356–369. Springer, 2010. 00160.

-
- [TTMI06] R Takaki, J Toriwaki, S Mizuno, and R Izuhara. Shape analysis of petroglyphs in central asia. *Forma*, 21:243–258, 2006.
- [TvW02] Alexandru Telea and Jarke J. van Wijk. An augmented fast marching method for computing skeletons and centerlines. In *Proc. of the Symposium on Data Visualisation 2002*, VISSYM '02, pages 251–ff, Aire-la-Ville, Switzerland, 2002. Eurographics Ass.
- [VF08] A. Vedaldi and B. Fulkerson. VLFeat: An open and portable library of computer vision algorithms, 2008.
- [Vin94] Luc Vincent. Morphological area openings and closings for grey-scale images. In *Shape in Picture*, pages 197–208. Springer, 1994.
- [VZ03] M. Varma and A. Zisserman. Texture classification: are filter banks necessary? In *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, pages II–691–8, Madison, WI, USA, 2003.
- [VZ05] Manik Varma and Andrew Zisserman. A statistical approach to texture classification from single images. *International Journal of Computer Vision*, 62:61–81, April 2005.
- [WFFD11] Marcus Wallenberg, Michael Felsberg, Per-Erik Forssén, and Babette Dellen. Channel coding for joint colour and depth segmentation. *Pattern Recognition*, 6835:306–315, 2011.
- [WHH03] Eric Wahl, Ulrich Hillenbrand, and Gerd Hirzinger. Surflet-pair-relation histograms: a statistical 3d-shape representation for rapid classification. In *Proceedings of Fourth International Conference on 3-D Digital Imaging and Modeling, 2003. 3DIM 2003.*, pages 474–481. IEEE, 2003. 00126.
- [Whi05] David S Whitley. *Introduction to rock art research*. Left Coast Press Walnut Creek, CA, 2005.
- [WSZ16] Ewald Wieser, Markus Seidl, and Matthias Zeppelzauer. A study on skeletonization of complex petroglyph shapes. *Multimedia Tools and Applications*, pages 1–19, 2016.
- [XWLB10] Yao Xu, Bo Wang, Wenyu Liu, and Xiang Bai. Skeleton graph matching based on critical points using path similarity. In *Computer Vision-ACCV 2009*, page 456–465. Springer, 2010.
- [YBYZ09] Xiaojun Yang, Xiang Bai, Xingwei Yang, and Luan Zeng. An efficient quick algorithm for computing stable skeletons. In *Proceedings of the 2nd International Congress on Image and Signal Processing (CISP'09)*, pages 1–5, October 2009.

REFERENCES

- [YKR08] Mingqiang Yang, Kidiyo Kpalma, and Joseph Ronsin. A survey of shape feature extraction techniques. *Pattern recognition*, page 43–90, 2008. 00163.
- [YLH⁺09] Xu-Cheng Yin, Qian Liu, Hong-Wei Hao, Zhi-Bin Wang, and Kaizhu Huang. A rock structure recognition system using FMI images. In *Proceedings of the 16th International Conference on Neural Information Processing: Part I, ICONIP '09*, page 838–845, Berlin, Heidelberg, 2009. Springer-Verlag.
- [ZBVH09] Andrei Zaharescu, Edmond Boyer, Kiran Varanasi, and Radu Horaud. Surface feature detection and description with applications to mesh matching. In *IEEE Conference on Computer Vision and Pattern Recognition, 2009. CVPR 2009.*, pages 373–380. IEEE, 2009.
- [ZL02] Dengsheng Zhang and Guojun Lu. Generic fourier descriptor for shape-based image retrieval. In *Proceedings of IEEE International Conference on Multimedia and Expo, 2002. ICME '02.*, volume 1, pages 425–428 vol.1, 2002.
- [ZL04] Dengsheng Zhang and Guojun Lu. Review of shape representation and description techniques. *Pattern Recognition*, 37(1):1–19, January 2004.
- [ZMLS07] Jianguo Zhang, Marcin Marszałek, Svetlana Lazebnik, and Cordelia Schmid. Local features and kernels for classification of texture and object categories: A comprehensive study. *International journal of computer vision*, 73(2):213–238, 2007.
- [ZPS⁺15] Matthias Zeppelzauer, Georg Poier, Markus Seidl, Christian Reinbacher, Christian Breiteneder, Horst Bischof, and Samuel Schuster. Interactive segmentation of rock-art in high-resolution 3d reconstructions. In *2015 Digital Heritage*, volume 2, pages 37–44, Sept 2015.
- [ZPS⁺ss] Matthias Zeppelzauer, Georg Poier, Markus Seidl, Christian Reinbacher, Samuel Schuster, Christian Breiteneder, and Horst Bischof. Interactive 3d segmentation of rock-art by enhanced depth maps and gradient preserving regularization. *ACM Journal on Computing and Cultural Heritage*, In Press.
- [ZS15] Matthias Zeppelzauer and Markus Seidl. Efficient image-space extraction and representation of 3d surface topography. In *IEEE International Conference on Image Processing (ICIP)*, pages 2845–2849, Sept 2015.
- [ZTKP14] L.V. Zotkina, A.S. Tekhterekov, V.M. Kharevich, and H. Plisson. An Experimental Study of Percussion Technologies in the Minusinsk

- Basin: Percussion and Tool Types¹. *Archaeology, Ethnology and Anthropology of Eurasia*, 42(1):55–65, March 2014.
- [ZTW⁺09] Zhiping Zeng, Anthony KH Tung, Jianyong Wang, Jianhua Feng, and Lizhu Zhou. Comparing stars: On approximating graph edit distance. *Proceedings of the VLDB Endowment*, 2(1):25–36, 2009.
- [ZWKL09] Qiang Zhu, Xiaoyue Wang, Eamonn Keogh, and Sang-Hee Lee. Augmenting the generalized Hough transform to enable the mining of petroglyphs. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '09, page 1057–1066, New York, NY, USA, 2009. ACM.
- [ZWKL10] Qiang Zhu, Xiaoyue Wang, Eamonn Keogh, and Sang-Hee Lee. An efficient and effective similarity measure to enable data mining of petroglyphs. *Data Mining and Knowledge Discovery*, 23(1):91–127, September 2010.
- [ZWKL11] Qiang Zhu, Xiaoyue Wang, Eamonn Keogh, and Sang-Hee Lee. An efficient and effective similarity measure to enable data mining of petroglyphs. *Data Mining and Knowledge Discovery*, 23(1):91–127, July 2011.
- [ZWS15] Matthias Zeppelzauer, Ewald Wieser, and Markus Seidl. A novel annotation tool for complex petroglyph shapes. In *The Future of Datasets in Vision Workshop (in conjunction with CVPR 2015)*, Boston, MA, USA, 06 2015.
- [ZZJS16] Matthias Zeppelzauer, Bartosz Zieliński, Mateusz Juda, and Markus Seidl. Topological descriptors for 3d surface analysis. In Alexandra Bac and Jean-Luc Mari, editors, *Computational Topology in Image Context: 6th International Workshop, CTIC 2016, Marseille, France, June 15-17, 2016, Proceedings*, pages 77–87. Springer International Publishing, 2016.

REFERENCES

Appendices

3D dataset

In this Appendix, we present an overview of our 3D dataset. Table A.1 contains a summary of all 3D scans. Table A.2 encloses basic measures of the scans. Figures A.1 to A.26 show orthophotos of all scans. The dataset has been acquired by the german 3D-surveying specialist ARCTRON¹ for the 3D-Pitoti project.

¹<http://www.arctron.de>

A. 3D DATASET

Table A.1: Summary of the reconstructed surfaces in the dataset. For each surface we provide an ID, description, its location (in terms of site, rock, and area) and the reconstruction method.

ID	Description	Site	Rock	Area	Reconstr. Method
1	Rosa Camuna	Foppe di Nadro	24	7	SLS
2	Warrior	Foppe di Nadro	24	7	SLS
3	Deer	Foppe di Nadro	24	9	SLS
4	Stele (part 1)	Naquane	Ossimo 8	11	SLS
5	Stele (part 2)	Naquane	Ossimo 8	11	SLS
6	Stele (part 3)	Naquane	Ossimo 8	11	SLS
7	Stele (part 4)	Naquane	Ossimo 8	11	SLS
8	Standing Rider (part 1)	Naquane	50	14	SFM
9	Standing Rider (part 2)	Naquane	50	14	SFM
10	Sun-Shape Superimposition	Naquane	50	15	SFM
11	Warrior Scene (part 1-1)	Seradina	12C	1	SLS
12	Warrior Scene (part 1-2)	Seradina	12C	1	SLS
13	Warrior Scene (part 1-3)	Seradina	12C	1	SLS
14	Warrior Scene (part 1-4)	Seradina	12C	1	SLS
15	Warrior Scene (part 2-2)	Seradina	12C	1	SLS
16	Warrior Scene (part 2-3)	Seradina	12C	1	SLS
17	Warrior Scene (part 2-4)	Seradina	12C	1	SLS
18	Warrior Scene (part 3-1)	Seradina	12C	1	SLS
19	Warrior Scene (part 3-2)	Seradina	12C	1	SLS
20	Warrior Scene (part 3-3)	Seradina	12C	1	SLS
21	Hunter With Bow	Seradina	12C	4	SLS
22	Hunter With Spear (part 1)	Seradina	12C	5	SLS
23	Hunter With Spear (part 2)	Seradina	12C	5	SLS
24	Hunting Scene (part 1)	Seradina	12C	10	SLS
25	Hunting Scene (part 2)	Seradina	12C	10	SLS
26	Hunting Scene (part 3)	Seradina	12C	10	SLS

Table A.2: Overview of basic measures of the digitized surfaces: the covered area (in pixels at 300dpi and in cm²), the number of 3D points in the point cloud, the percentage of pecked regions, the number of disconnected pecked regions and the range of depth values.

ID	Covered Area		Num.	Percentage	Num.	Depth Range
	in px	in cm ²	3D Pts.	Pecked	Seg.	in mm
1	5 143 296	368.69	3 264 005	14.61	48	2.89
2	15 638 394	1121.03	10 280 976	10.56	21	4.83
3	8 846 214	634.14	5 503 742	47.63	18	9.11
4	15 507 622	1 111.66	3 782 381	14.96	17	62.52
5	16 994 561	1 218.25	2 658 330	17.27	44	70.60
6	13 102 254	939.23	1 260 401	12.67	13	49.32
7	12 035 386	862.75	810 312	34.02	26	15.17
8	12 834 446	920.03	8 677 163	26.17	45	6.74
9	12 835 586	920.11	8 386 259	32.83	29	3.82
10	5 901 454	423.04	2 096 476	21.59	9	5.41
11	5 632 144	403.74	3 541 799	9.26	23	10.23
12	7 103 936	509.24	4 432 013	5.09	6	10.22
13	6 155 628	441.26	3 810 000	8.26	63	19.85
14	5 855 280	419.73	4 417 779	6.47	17	10.50
15	4 855 764	348.08	2 981 570	4.44	24	9.39
16	4 029 231	288.83	2 523 543	6.58	29	4.27
17	4 838 487	346.84	3 022 433	3.15	27	21.75
18	6 396 152	458.50	4 007 232	19.41	25	9.45
19	7 141 253	511.92	4 472 845	18.20	32	17.32
20	6 864 476	492.08	4 238 990	12.02	15	21.39
21	3 909 579	280.26	2 255 030	20.40	61	5.32
22	4 073 804	292.03	2 395 125	16.34	65	3.99
23	3 612 131	258.93	2 113 670	24.23	54	5.33
24	19 104 798	1 369.52	10 685 564	26.61	152	27.35
25	14 920 005	1 069.53	8 188 025	15.55	63	17.49
26	8 921 684	639.55	5 515 973	15.59	99	16.62
Overall	232 253 565	16 648.97	115 321 636	18.68	1 025	70.60

A. 3D DATASET



Figure A.1: Orthophoto of scan ID 1.



Figure A.2: Orthophoto of scan ID 2.



Figure A.3: Orthophoto of scan ID 3.



Figure A.4: Orthophoto of scan ID 4.

A. 3D DATASET



Figure A.5: Orthophoto of scan ID 5.



Figure A.6: Orthophoto of scan ID 6.



Figure A.7: Orthophoto of scan ID 7.



Figure A.8: Orthophoto of scan ID 8.



Figure A.9: Orthophoto of scan ID 9.



Figure A.10: Orthophoto of scan ID 10.



Figure A.11: Orthophoto of scan ID 11.



Figure A.12: Orthophoto of scan ID 12.



Figure A.13: Orthophoto of scan ID 13.



Figure A.14: Orthophoto of scan ID 14.

A. 3D DATASET



Figure A.15: Orthophoto of scan ID 15.



Figure A.16: Orthophoto of scan ID 16.



Figure A.17: Orthophoto of scan ID 17.



Figure A.18: Orthophoto of scan ID 18.



Figure A.19: Orthophoto of scan ID 19.



Figure A.20: Orthophoto of scan ID 20.

A. 3D DATASET



Figure A.21: Orthophoto of scan ID 21.



Figure A.22: Orthophoto of scan ID 22.



Figure A.23: Orthophoto of scan ID 23.

A. 3D DATASET



Figure A.24: Orthophoto of scan ID 24.



Figure A.25: Orthophoto of scan ID 25.



Figure A.26: Orthophoto of scan ID 26.

Peer-reviewed Publications

The following peer-reviewed publications contain results of the thesis:

Journal Articles

Markus Seidl, Ewald Wieser, and Craig Alexander. Automated classification of petroglyphs. *Digital Applications in Archaeology and Cultural Heritage*, 2(2,3):196 – 212, 2015

Ewald Wieser, Markus Seidl, and Matthias Zeppelzauer. A study on skeletonization of complex petroglyph shapes. *Multimedia Tools and Applications*, pages 1–19, 2016

Matthias Zeppelzauer, Georg Poier, Markus Seidl, Christian Reinbacher, Samuel Schuler, Christian Breiteneder, and Horst Bischof. Interactive 3d segmentation of rock-art by enhanced depth maps and gradient preserving regularization. *ACM Journal on Computing and Cultural Heritage*, In Press

Conference Papers

Markus Seidl and Christian Breiteneder. Detection and classification of petroglyphs in gigapixel images – preliminary results. *VAST11: The 12th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage - Short Papers*, pages 45–48, 2011

Markus Seidl and Christian Breiteneder. Automated petroglyph image segmentation with interactive classifier fusion. In *Proceedings of the Eighth Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP '12)*, 2012

Matthias Zeppelzauer and Markus Seidl. Efficient image-space extraction and representation of 3d surface topography. In *IEEE International Conference on Image Processing (ICIP)*, pages 2845–2849, Sept 2015

Matthias Zeppelzauer, Georg Poier, Markus Seidl, Christian Reinbacher, Christian Breiteneder, Horst Bischof, and Samuel Schuler. Interactive segmentation of rock-art in high-resolution 3d reconstructions. In *2015 Digital Heritage*, volume 2, pages 37–44, Sept 2015 - *Best Paper Award*

B. PEER-REVIEWED PUBLICATIONS

Workshop Papers

Markus Seidl, Ewald Wieser, Matthias Zeppelzauer, Axel Pinz, and Christian Breiteneder. Graph-Based Shape Similarity of Petroglyphs. In Lourdes Agapito, Michael M. Bronstein, and Carsten Rother, editors, *Computer Vision - ECCV 2014 Workshops*, volume 8925 of *Lecture Notes in Computer Science*, pages 133–148. Springer International Publishing, 2015

Matthias Zeppelzauer, Bartosz Zieliński, Mateusz Juda, and Markus Seidl. Topological descriptors for 3d surface analysis. In Alexandra Bac and Jean-Luc Mari, editors, *Computational Topology in Image Context: 6th International Workshop, CTIC 2016, Marseille, France, June 15-17, 2016, Proceedings*, pages 77–87. Springer International Publishing, 2016

Poster

Matthias Zeppelzauer, Ewald Wieser, and Markus Seidl. A novel annotation tool for complex petroglyph shapes. In *The Future of Datasets in Vision Workshop (in conjunction with CVPR 2015)*, Boston, MA, USA, 06 2015