

FRAMEWORK FOR AUTOMATIC IDENTIFICATION OF PAPER WATERMARKS  
WITH CHAIN CODES

A DISSERTATION  
in  
Electrical and Computer Engineering  
and  
Telecommunications and Computer Networking

Presented to the Faculty of the University of  
Missouri - Kansas City in partial fulfillment of  
the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

PLAMEN DOYNOV

Dipl. Eng., Technical University of Sofia, Sofia, Bulgaria  
M.E.E., Catholic University of America, Washington, D.C., USA

Kansas City, Missouri  
2017

© 2017

PLAMEN DOYNOV

ALL RIGHTS RESERVED

FRAMEWORK FOR AUTOMATIC IDENTIFICATION OF PAPER WATERMARKS  
WITH CHAIN CODES

Plamen Doynov, Candidate for the Doctor of Philosophy

University of Missouri-Kansas City, 2017

ABSTRACT

In this dissertation, I present a new framework for automated description, archiving, and identification of paper watermarks found in historical documents and manuscripts. The early manufacturers of paper have introduced the embedding of identifying marks and patterns as a sign of a distinct origin and perhaps as a signature of quality. Thousands of watermarks have been studied, classified, and archived. Most of the classification categories are based on image similarity and are searchable based on a set of defined contextual descriptors. The novel method presented here is for automatic classification, identification (matching) and retrieval of watermark images based on chain code descriptors (CC). The approach for generation of unique CC includes a novel image preprocessing method to provide a solution for rotation and scale invariant representation of watermarks. The unique codes are truly reversible, providing high ratio lossless compression, fast searching, and image matching. The development of a novel distance measure for CC comparison is also presented. Examples for the complete process are given using the recently acquired watermarks digitized with hyper-spectral imaging of *Summa Theologica*, the work of Antonino Pierozzi (1389 – 1459). The performance of the algorithm on large datasets is demonstrated using watermarks datasets from well-known library catalogue collections.

## APPROVAL PAGE

The faculty listed below, appointed by the Dean of the School of Graduate Studies, have examined a thesis titled "Framework for Automatic Identification of Paper Watermarks with Chain Codes," presented by Plamen Doynov, candidate for the Doctor of Philosophy degree, and certify that in their opinion it is worthy of acceptance.

### Supervisory Committee

Reza Derakhshani, Ph.D., Committee Chair  
Department of Computer Science and Electrical Engineering

Ghulam M. Chaudhry, Ph.D.  
Department of Computer Science and Electrical Engineering

Appie Van de Liefvoort, Ph.D.  
Department of Computer Science and Electrical Engineering

Deendayal Dinakarandian, Ph.D.  
Department of Computer Science and Electrical Engineering

Jeffrey Rydberg-Cox, Ph.D.  
Department of English, Director of Classical and Ancient Studies Program,  
Affiliated Faculty, Department of Computer Science and Electrical Engineering

## CONTENTS

ABSTRACT .....	iii
LIST OF ILLUSTRATIONS .....	ix
LIST OF TABLES .....	xiv
ACKNOWLEDGMENTS .....	xv
Chapter	
1. INTRODUCTION .....	1
1.1 Research motivation.....	2
1.2 Thesis objective .....	3
1.3 Thesis contributions .....	5
1.4 Thesis overview .....	6
2. PAPER AND PAPER WATERMARKS .....	9
2.1 The history of paper .....	9
2.2 Origin and evolution of paper watermarks .....	13
2.3 The Study of watermarks—literature review.....	19
2.3.1 Imaging and reproduction techniques for watermarks .....	25
2.3.2 Description, classification and storage .....	29
2.3.3 Current methods for search and retrieval .....	31
2.4 Hyperspectral imaging of historical documents.....	52
2.4.1 Hyperspectral imaging system for historical documents .....	60
2.4.2 Processing of the hyperspectral cube .....	75
2.4.3 Watermark image extraction .....	87
2.5 Watermark image preprocessing.....	88

2.6	Watermark image binarization and skeletonization.....	92
2.7	An alternative non-contact method for watermarks reproduction .....	97
2.8	Conclusions.....	100
3.	AUTOMATIC IDENTIFICATION OF PAPER WATERMARKS .....	102
3.1	Introduction.....	102
3.2	Shape descriptors .....	102
3.3	Chain codes as image descriptors .....	111
3.4	Chain code generation for watermarks (from image to chain code).....	115
3.5	Reverse transform (from chain code to the original image) .....	118
3.6	Chain code as a lossless compression.....	119
3.7	Limitations of chain codes .....	126
3.8	Conclusions.....	126
4.	ROTATION, SCALE AND TRANSLATION INVARIANT CHAIN CODE .....	129
4.1	Introduction.....	129
4.2	Watermark minimum boundary box (MBB) .....	130
4.3	Minimum boundary box attributes.....	132
4.4	Minimum boundary box for rotation and scale invariance .....	133
4.5	Watermark recto and vero images .....	134
4.6	Conclusions.....	137
5.	COMPARISON OF RST_INVARIANT CHAIN CODES .....	138
5.1	Introduction.....	138
5.2	Comparing chain codes with the same length.....	139
5.3	Comparing chain codes with different length.....	140

5.4	Chain code distance (CCD) as measure of watermark similiarity .....	142
5.5	Application of CCD for watermarks comparison .....	146
5.6	Conclusions.....	153
6.	<b>AUTOMATIC IDENTIFICATION OF WATERMARKS WITH CHAIN CODES .....</b>	<b>155</b>
6.1	Introduction.....	155
6.2	Watermark datasets .....	155
6.3	Performance evaluation .....	156
6.3.1	Dataset sample size .....	157
6.3.2	receiver operating characteristics.....	159
6.3.3	Experimental results using CCD.....	160
6.4	Alternative methods for watermark identifications .....	164
6.4.1	Machine learning and image matching.....	165
6.4.2	Convolutional Neural Networks (CNN) .....	166
6.4.3	Watermarks dataset for use with CNN .....	167
6.4.4	Experimental results of watermarks matching using CNN.....	175
6.4.5	Bag of visual words .....	178
6.4.6	Experimental results using BoVW.....	176
6.5	Conclusions.....	182
7.	<b>WATERMARK COMPOSITE FEATURE VECTOR.....</b>	<b>183</b>
7.1	Introduction.....	183
7.2	Watermarks complexity .....	183
7.3	Extended chain code array .....	184
7.4	Structure of composite feature vector (CFV).....	185

7.5	Composition and use of CFV.....	187
7.6	Examples with CFV application .....	188
7.7	Conclusions.....	192
8.	SUMMARY .....	194
8.1	Discussion.....	194
8.2	Conclusions.....	195
8.3	Future work.....	196
Appendix		
A.	WATERMARKS FROM THE BERNSTEIN COLLECTION USED IN THIS STUDY.....	199
B.	THE ORIGINAL AND TRANSFORMED IMAGES OF WATERMARKS .....	206
C.	THE TRANSFORMED AND SCALED IMAGES OF WATERMARKS .....	216
D.	EXAMPLE OF CHAIN CODE .....	218
BIBLIOGRAPHY.....		220
VITA.....		236



## LIST OF ILLUSTRATIONS

Figure	Page
1.1: Digital images of the oldest known paper watermarks. Greek cross (#5410) and name of paper maker (#12021). C. Briquet, publishes <i>Les filigranes</i> .....	1
1.2: Prima, Secunda and Tertia of <i>Summa Theologica</i> the work of Antonino Pierozzi (1389 – 1459).....	2
2.1: Partners in the Bernstein Consortium. ....	11
2.2: The Memory of Paper integrated environment.....	12
2.3: Early production of paper in China during Han dynasty and Europe.....	14
2.4: An example of a wire figure added to the chain and laid mould wires .....	14
2.5: Hand mould with relatively complex watermark. The verge and warp wire are visible attaching the wire figure to the chain and laid wires.....	15
2.6 : “Pair of twins” watermarks, c. 1640, Swedish National Archive.....	24
2.7 : Different states of the same watermark. Swedish National archive, c 1825. The “F” has not changed while the “C” has shifted and changed shape noticeably.....	25
2.8: Examples of watermarks digitized with different techniques and respectively different quality. ....	28
2.9: Bull’s head twin pairs, <i>Missale Speciale</i> , c. 1473. Two pairs from left to right: Wide face - Square face and Level eye - Slant eye. Adapted from.....	30
2.10: Examples for description of simple and complex watermarks. IPH Standard, Version 2.1.1, 2013.....	32
2.11: Lion class with subclasses. IPH Standard, Version 2.1.1, 2013. ....	33
2.12: For a simple search the user just enters the search terms in a search field .....	40
2.13: For an advanced search the user can combine several search fields.....	40
2.14: The browsing with a motif offers the possibility to navigate in the tree structure of the systematics classification by names or by icons .....	40
2.15: For a simple search users enters a search term in the Keyword Search field. ....	41
2.16: For an advanced search up to 8 parameters may be used:1. 41Creator/Artist; 2. Signatures; 3. Time Period; 4. Keyword; 5. Range; 6. Exact value; 7. Not Empty Value; 8. Empty Value.....	41
2.17: Bernstein Paper Studies Kit – links, software and manuals for the studies of watermarks. 42	
2.18: Internet page to access watermarks by using the class as the retrieval mechanism. (a) Textual list of classes. (b) Iconic representation of the classes. (c) Main page for the class Eagle (Aigle in French.....	45

2.19: Internet page to access watermarks by using the class as the retrieval mechanism. (a) Textual list of classes. (b) Iconic representation of the classes. (c) Main page for the class Eagle (Aigle in French) .....	45
2.20: (a) Watermark Main page. (b) Page that allows the setting of the different values. (c) Result page returned by a query .....	46
2.21: Global features on watermarks. (a) Space between chain lines and the density of laid lines per unit length; (b) Height and width of the watermark; (c) Number of regions; (d) Number and position of junctions; (e) Position of the watermark depending on the chain lines .....	46
2.22: Graphical representation of the circular histogram. (a) The center of gravity of the watermark is used as the middle point; (b) The number of pixels present inside a radial segment. ....	47
2.23: Eight filters $F_j$ used to compute the directional information on the images.....	47
2.24: (a) Main page of the similarity retrieval module; (b) Drawing applet interface to draw a similar shape; (c) Result of similar images .....	49
2.25: (a) Main page of the similarity retrieval module; (b) Drawing applet interface to draw a similar shape; (c) Result of similar images. ....	49
2.26: Claude Monet's A Haystack in the Evening Sun (c 1891), Courtesy of Gosta Serlachius, Fine Arts Foundation was examined at the University in the Recenart – Research Centre for Art project.....	54
2.27: Ocean Optics Jaz EL350 VIS-NIR spectrometer (left) and LabVIEW based GUI for control and monitoring .....	61
2.29: Performance of the integrated Jaz and VariSpec modules .....	646
2:30: Control panel for VariSpec and Jaz with the spectra of 3200K light source and the overhead room fluorescent light on .....	67
2:31: Spectra reader for binary and concatenated binary spectral records .....	66
2.32: The Canon EOS T5i camera with the VariSpec VIS LCT filter attached. ....	69
2.33a: Optical performance test at 520 nm wavelength.....	70
2.33b: Optical performance test at 560 nm wavelength. ....	70
2.33c: Optical performance test at 560 nm wavelength.....	69
2.34: CODICES imaging of <i>Summa Theologica</i> .....	72
2.35: The RGB jpeg image of <i>Summa Theologica</i> , Volume 1, page 1.....	73
2.36: The RGB jpeg image of <i>Summa Theologica</i> , Volume 1, page 1 .... 712.37: Grayscale images from 400nm to 700nm and the calculated relative SNR .....	73
2.37: Grayscale images from 400nm to 700nm and the calculated relative SNR .....	74
2.38: Spectrogram of “white” flat light source .....	75
2.39: Two examples of typical hyperspectral cubes .....	76

2.40: GUI for interactive conversion of CR2 files to TIFF, PNG, BMP, and other image formats .....	77
2.41: Individual and joint image histograms for three wavelengths .....	78
2.42: PCA applied to hyperspectral images of <i>Summa Theologica</i> (top row). The first three principal components are shown next to the RGB image of the ROI. ....	80
2.43: ICA algorithm outputs. Images from Adair Chant Book.....	84
2.44: RANSAC alignment of images – point correspondences.....	86
2.45: Two images (A and B), the error of RANSAC alignment (C), the warped image by features based alignments with RANSAC (D), warped image by features based alignments with Least-squares (E), , and error of Least-squares alignment (F). ....	86
2.46: Images alignment with SIFT transform: images (A and B), SIFT field flow (C), and the registration error (D).....	87
2.47: Blue layer of a <i>Summa Theologica</i> RGB recto image (A) and ROI (B). Global thresholding (C) performs worse than Local thresholding (D). ....	90
2.48: Blue layer of a <i>Summa Theologica</i> RGB recto image (A) and ROI (B). Global thresholding (C) performs worse than Local thresholding (D).. ....	90
2.49: Interactive digital restoration with high degree polynomial curves .....	91
2.50: Image with watermark from Volume III of <i>Summa Theologica</i> (left) and the binary image of the extracted watermark. ....	92
2.51: Binary image of a rectangle (A), and the result from the skeletonization (B) and Medial axis transform (C).....	92
2.52: Examples for two possible patterns and the generated values A and B using the functions $F_a$ and $F_b$ .....	96
2.53: Lillie watermark original image, its binary image and its skeleton .....	97
2.54: LED backlight with the preprocessed image of <i>Fleur a six petals</i> watermarks (left). The tracing is performed with black marker (right). ....	98
2.55: Briquet's <i>Fleur a six petals</i> watermark (A). A six-petal flower image from <i>Summa Theologica</i> after processing (B). The same watermark traced from the image (C). The watermark binary image (D) and its skeleton image (E).....	98
2.56: The preprocessed <i>Letter R</i> watermark image (1A) and its inverse (2A). The two images after trace reproduction (1B & 2B), binarization (1C & 2C), and after skeletonization (1D & 2D).....	99
3.1: Description methods for shapes and contours .....	103
3.2: Chain code generation schema with four and eight neighboring pixel.....	113
3.3: CC generation for two watermarks after binary and skeleton transforming of the original images .....	116
3.4: The Chain codes of the four watermarks are very different .....	117

3.5: Watermark restored from their CC (top row) and displayed with yellow color on the bottom row for comparison with the original watermarks. Differences are in red. ....	116
3.6: Comparing the size of the original watermark jpg image and its chain code.....	121
3.7: Comparing the size of the watermark's CC with different representations.....	122
3.8: CC for Bernstein watermark #004001001 expressed as 585 decimal numeric characters and its ASCII base 85 representation with 280 characters.....	124
4.1: Capturing WM relative orientation and scale features (see text).....	130
4.2: Comparison of superimposed rectangles. The red rectangles are based on major axis alignment (A). The yellow rectangle is the minimum boundary box drawn by the algorithm of the new method (B). ....	132
4.3: Watermarks from Bernstein collection. Are they four separate watermarks or <i>recto</i> and <i>verso</i> images of two watermarks? .....	134
4.4: Binary <i>recto</i> and <i>verso</i> images of Crown watermark with the location of their respective centroids. ....	135
4.5: Centroids of the watermark skeleton images indicating <i>recto</i> and <i>verso</i> identity .....	136
5.1: Distance metrics for CC compared to themselves. The calculated differences have 0 values. The similarity score is the sum of the values on the diagonal.....	140
5.2: New Distance metric for CC comparison. The similarity score of 598 is the sum of the values on the diagonal.....	141
5.3: Supremum and infimum of two sets A and B.....	145
5.4: Comparison of watermarks based on their chain codes and chain codes distances. The CC differences for watermark <i>Anker</i> and watermark <i>Hand</i> with similarity score 640. The difference between watermark <i>Bell</i> and watermark <i>Hand</i> is 387. ....	147
5.5: Self-comparison for the watermark <i>Anker</i> (left) and self-comparison for the watermark <i>Bell</i> (right). As expected the difference measure is zero indicating a perfect match .....	148
5.6: Chain codes of Crown watermark <i>recto</i> (left) and <i>verso</i> (right). The Chain codes have the same length and similar patterns .....	148
5.7: Crown watermark <i>recto</i> and <i>verso</i> comparison. Self-comparison CCD is 0 (left). The <i>recto vs. verso</i> CCD is 596.....	149
5.8: Watermark images of class Anchor .....	150
5.9: Chain codes distances for class Anchor watermarks. Two watermark images are copied and added to the original dataset for testing. They have formed two matching pairs (3 & 6) and (7 & 11) .....	151
5.10: The watermarks from Bernstein collection from Figure 4.3 with their mirror images. ....	152
5.11: CCD for the four watermarks and their mirror images.....	152

6.1: Distance between the CC of 250 WM with pair-wise comparison.....	161
6.2: ROC based on CCD for the degree of degradations given in Table 6.1 .....	162
6.3: The direction probability of CC formation. Due to the skeletonization, at pixel $P_i$ , coming from pixel $P_{i-1}$ , the direction of moving forward with higher probability is indicated with darker blue.....	163
6.4: Synthetic image distortions. From left to right top row: original checkerboard image, rotation, affine, reflective and non-reflective similarity. Bottom row: projective, polynomial, piece-wise linear, barrel and pin-cushion transforms .....	168
6.5: WM images used in CNNexample .....	173
6.6: Examples of transformed images for CNN training .....	174
6.7: CNN train with 50 epochs and performance test results (insert) for 100 x 100 size images.....	175
6.8: An example of c8/p2/c16/p2 CNN architecture .....	176
6.9: Mean Square Error for c8/p2/c16/p2 CNN.....	177
6.10: Classification accuracy of 48.13 % for 10 watermarks using limited training with 150 randomly selected watermark images.....	180
6.11: Classification accuracy of 89.38 % with 30 images per class selected randomly for training (300 out of 400) and a disjointed set of 10 images for testing.....	181
6.12: Classification accuracy of 96.88 % with 39 images per class selected randomly for training (390 out of 400) and the rest for testing.....	181
7.1: Examples of watermark images requiring rigorous preprocessing or with complex motives .....	184
7.2: Watermarks of class <i>Bell</i> . From a simple figure to one with two additional elements.....	189
7.3: Example of image processing for generation of CCs and formation of the CFV .....	190
7.4: Recognition of similar watermarks with additional motives using CFV.....	192

## LIST OF TABLES

Table	Page
2.1: Different implementations of ICA algorithms.....	83
2.2: Rigid, Non-rigid and Optical flow alignment algorithms.....	85
3.1: Example of CC, DCC, and shape number for a simple contour.....	114
3.2: Chain code for the Bernstein #004001001 watermark (585 bytes).....	122
3.3: Chain code in binary format representation (220 bytes).....	122
3.4:Chain code forming the the first 24-bit group.....	124
6.1: EER, AUC, and threshold for different degrees of degradation.....	162

## ACKNOWLEDGMENTS

This dissertation is a result of the encouragement and support of many people who made me the person I am, people that I remember and I admire.

First, many thanks to my teachers who through all these years gave me knowledge and further inspired my curiosity to explore the world. I would like to thank my Ph.D. research advisor, Dr. Reza Derakhshani, for all his support over the years in helping me to become a better scientist and engineer. He always encouraged me to face new challenges and continuously helped me in improving my skills. Dr. Derakhshani inspired me with his grasp on modern technologies and persistent approach to research and laboratory experimentation. His interdisciplinary interests made it possible for me to be part of the CODICES team of UMKC and apply my skills to the forensic exploration of historical documents and manuscripts.

I would also like to thank all my Ph.D. committee members for being supportive throughout my graduate career. I am grateful to Dr. Ghulam Chaudhry for always being available to share knowledge and wisdom and for giving me opportunity to be his teaching assistant. I am thankful to Dr. Van de Liefvoort (Apee) for all his support and guidance. Thank you for coming to all my presentations and to the meetings of IEEE. Thank you for trusting me to teach several classes. I am also thankful to Dr. Deendayal Dinakarbandian (Dinakar) for being a valuable and supportive member of my Committee. Thank you for always being responsive to meet and discuss scientific and engineering problems from an interdisciplinary point of view.

I express my appreciations to Dr. Rydberg-Cox for introducing me to the enchanting world of paper making and filigranology, the study of paper watermarks. Thank you for becoming a

member of my Ph.D. Committee. Your knowledge of classical and ancient documents and your broader understanding of the underlying problems has been an invaluable guidance in my interdisciplinary research.

I am thankful to Dr. Pavan Tankasala for many years of great collaboration in the Computational Intelligence and Bio-Identification Lab of UMKC. I am also thankful to Dr. Nathan Oyler for many discussions and collaboration during the development of the hyperspectral imaging platforms. And Dr. Virginia Blanton for her support and collaboration as part of the CODICES team.

And of course, I thank my colleagues and friends at MRIGlobal, the company where I have been for almost twenty years and where the scientific challenges are the norm and a natural extension of my academic research. Special thanks to my collaborator and friend Dr. Paul Haney for his continuous support with many scientific endeavors.

Finally, and most importantly, I am deeply grateful to my family, especially my mother and father, for giving me life and a chance to be all that I want to be. My children have been and continue to be a great inspiration: Felice, Anya, Victor, and Sofia, thank you. “My dad can do anything...” is a very high standard to live with. Last, but not least, I would like to thank my wife, Nina Doynov, for her love and immense support.



*I dedicate this thesis*

*to the memory of my parents and to my family*

*for their continuous support and inspirational love.*

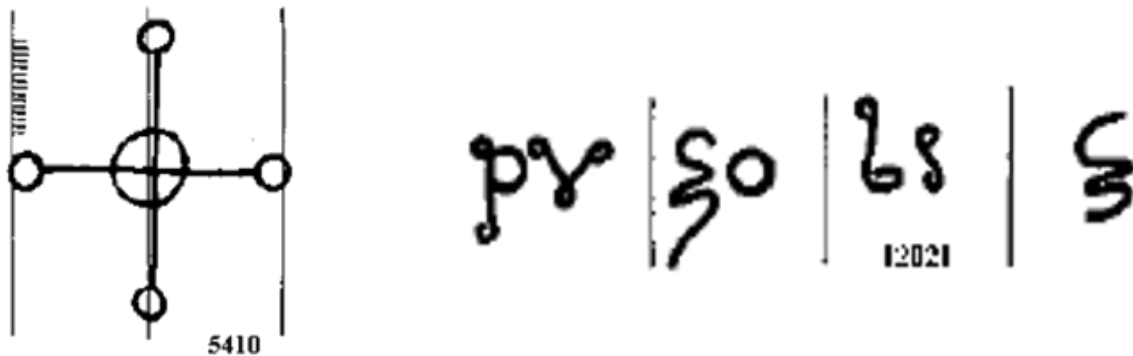
*I love you to the stars and back, infinity times...*

*P.D.*

## CHAPTER 1

### INTRODUCTION

Paper watermarks are invaluable to the scholars and enigmatic to the novice. Watermarks are interesting, beautiful and revealing. Intrigued by watermarks, people have been searching, discovering and cataloguing them for centuries [1-9]. The oldest known watermark was discovered by Briquet printed on paper produced in 1292 in Fabriano, Bologna, Italy [8]. Catalogued as watermark #5410, it represents a Greek cross with circles at the cross-center and cross-ends [Figure 1.1].



**Figure 1.1: Digital images of the oldest known paper watermarks. Greek cross (#5410) and name of paper maker (#12021). C. Briquet, publishes *Les filigranes* [8]**

Perhaps the very first watermarks were produced by broken mould wires leaving noticeable marks in the dried paper in addition to the inherent laid and chain lines artefact [5]. Soon after the paper production began in Europe, some manufacturers added wires to form simple shapes to leave an impression with a desired pattern on the finished paper product. The resulting effect was an embedded “wire-mark” left during the pulp’s drying process and historically have become known as “watermark”. Originally, watermarks may have served to identify papers produced by different workmen within a factory to mark specifically their production for which they were “paid by piece” [2]. They also were used as a proof of the manufacturing date, as an

indication of paper size and orientation during the paper sheet cutting process. Later on, over the centuries and nowadays, paper watermarks are used to identify paper owners, authenticity and also for protection from forgery and theft of important documents such as bank notes, passports, and others.

### 1.1 Research motivation

I was introduced to the world of historical documents and paper watermarks in 2014 when I joined CODICES, a multidisciplinary team of professors and graduate students at University of Missouri – Kansas City. The team was engaged in studying historical books and development of methods for digital imaging and artefacts discovery. In particular, I was involved in the hyperspectral imaging of *Summa Theologica* (Figure 1.2). The four-volume set is the work of Antonino Pierozzi (1389 – 1459) and contains many of the contemporary watermarks.



**Figure 1.2. Prima, Secunda and Tertia of *Summa Theologica* – the work of Antonino Pierozzi (1389 – 1459).**

In the process of learning about watermarks, their origin and historical importance, I also understood that there are no easy and automated methods for watermarks description, cataloging, and identification. Based on my experience in digital image processing and biometrics, I started to search for an applicable method to describe the watermarks in a concise

and unique mode. After all, the watermarks are a specific class (or classes) of shapes and there are numerous methods for shape description. In summary, I have identified an existing problem and I challenged myself to find a good and possibly an elegant solution.

## **1.2 Thesis objective**

The objective of this thesis is to develop a method for deterministic description of paper watermarks found in digital images of historical documents and manuscripts. The goal is to be able to describe a watermark in a way that can result in a unique and concise code that can be used for identification. The ability to generate such an ID would result in capability to facilitate the automatic classification, comparison, search and retrieval of watermark databases.

It turns out that the identified problem with watermarks description and classification is a persistent one. Currently, the description, classification, storage, search and retrieval of watermarks is text based. This approach is known as a content based image retrieval (CBIR) – a predefined list of categories and features used to describe a watermark for classification and identification purpose. Some of the difficulties of this approach arise from the different level of adherence to the existing standards for watermarks database development. Many datasets and collections in libraries are organized without an overarching and an interoperability governing standard. A set of predefined parameters and criteria are used for organizing the digital images of watermarks (qualitative type classification). The modern tools for data storage, organization, and online access have facilitated the use of the existing databases and their expansions. However, the process model still remains a highly intense work based on textual descriptions and search interactions. The development of better methods for archiving and search models is ongoing.

In their book “The grand design” [10], S. Hawkins and L. Mlodinow outline four important characteristics of a good model. They specify that a good model is:

1. *Elegant.*
2. *Contains few arbitrary or adjustable elements.*
3. *Agrees with and explains all existing observations.*
4. *Makes detailed predictions about future observations that can disprove or falsify the model if they are not borne out.*

The authors elaborate that it is not clear how to define the first characteristic: “elegant”. They referred to the Einstein’s famous quote “All thing have to be made simple, but not simpler than that.” Nevertheless, elegant and the other three characteristics are very applicable to the features of a good descriptor. Actually, a descriptor is a product of a method that follows a selected representation model. In this regard, ideally, the watermark descriptor would have the following properties:

- Universally applicable
- Easy to generate
- Unique (ability to capture watermark shapes and details)
- Complete (there is no need for any additional information)
- Quantitatively comparable (ability to measure the degree of watermarks’ similarity)

This thesis outlines the framework of a method for generation of unique, numerical descriptor for identification of digitized watermarks with the desired properties listed above.

### 1.3 Thesis contributions

The novel contributions in this thesis, titled “Framework for Automatic Identification of Paper Watermarks with Chain Codes”, are as follows:

1. Generation of unique numerical watermark descriptor based on a chain code.
2. Method for rotation, scale, and translation invariance of the generated chain codes (RST-invariant chain code).
3. Method for an extended *recto* and *verso* variability capture.
4. Method to compare two chain codes with different size – distance metric for similarity measurements.
5. Exact re-generative model based on chain code with very high ratio lossless compression of watermark images.
6. An alternative, new non-contact watermarks reproduction method.
7. Development of a composite feature vector (CFV) for complete description and comparison of watermarks.

All of the above contributions could be summarized as:

- ❖ Introduction of the C-gram as a unique descriptor of watermarks.

The design and implementation of the novel methods are described in the following chapters. The term “original” could be added to the proposed methods to indicate that they are introduced for the first time and are not an alternative (“new”) implementation of an existing method. In this regard, five of the seven contributions are original ideas and novel solutions with practical applications to paper watermarks. Additional supporting materials associated with generation of chain codes are also provided. Performance evaluation of the proposed methods is provided and also compared with a leading modern machine learning algorithms based on convolutional neural networks and bag of visual words. In the last chapter, after a summary and discussion of the proposed framework, I outline directions for further research.

## 1.4 Thesis overview

The previous sections in Chapter 1 have given an introduction to the importance of paper watermarks, the problems associated with studying them, and the motivation to develop a method for deterministically describing watermarks with ability to compare and identify them. The rest of this thesis is organized as follows:

**Chapter 2: Paper and paper watermarks** is a brief overview of the history of papermaking. It also describes the natural occurrence of the watermarks in the process of papermaking and the importance of studying watermarks. The literature review presents a coverage of background and literature surveys relevant to the research. The scope was focused on paper watermarks and their history, an introduction to the history of paper making and the origin and evolution of watermark creation. The chapter offers a review of the techniques used for capturing watermarks from historical documents. The current methods for description, classification, storage and retrieval of watermarks are outlined along with their limitations. The design of a modern hyperspectral imaging system is introduced with demonstration of experimental results. Some techniques for watermark images processing and restoration are presented. A new method for non-contact watermark reproduction is introduced and application examples are demonstrated.

**Chapter 3: Automated identification of paper watermarks** provides an overview of existing shape descriptors and justifies the applicability of chain code for watermark representation. The motivation for a new method of watermark description is further elaborated and the implementation is demonstrated. The generation of the chain code is presented with the reverse transform demonstrating that chain codes are inherently a re-generative lossless

compression engine. The compression of the chain code is extended with a novel 24-bit group presentation with ASCII base 85. The main limitations of the chain codes are described: 1) their rotation and scale dependence; and 2) non-existing ability to numerically compare them.

**Chapter 4: Rotation, scale and translation invariant chain code** demonstrates an original framework for generation of chain codes based on watermark signatures that are independent of the size (scale) and orientation (translation and rotation) of the watermark image. The minimum boundary box is introduced as a preprocessing technique that provides normalization of the watermark images in a dataset and is the basis for generation of rotation, scale and translation invariant (RST-invariant) chain codes. The *recto* and *verso* imaging variability of watermarks is also discussed and a solution for testing and invariant representation is proposed.

**Chapter 5: Comparison of chain code** introduces an original method for comparison of chain codes. The chapter elaborates that while the dependence of chain codes on rotation and scale have prevented their use, the ability to compare chain codes was also a dominant limiting factor. The review of research works on chain codes and the search for solutions to overcome their limitations are presented. Next, the new distance measure algorithm is presented with demonstration of its ability to compare chain codes with the same or different lengths. The application of the proposed chain code distance metric for watermarks' similarity measure is demonstrated.

**Chapter 6: Automatic identification of watermarks with chain codes** is presenting the results of the application of the new method to a real dataset of watermarks. The selected dataset is described with a statistical justification of the applicable sample test size. The performance of the new method is evaluated with receiver operating characteristics curves. For



the purpose, a random artificial degradation of the chain codes is introduced. To further demonstrate the level of performance of the new method, it is compared with the application of modern deep learning technique. A Convolutional neural network (CNN) was constructed and applied to the same set of watermark images. In order to expand the sample size, synthetic image degradation techniques are used. The expanded dataset was used with k-fold analysis and random separation of the data on training and testing subsets. The bag of visual words was also tested with ten classes of the same dataset and the classification results are presented with confusion matrices for different training and testing scenarios.

**Chapter 7: Watermark composite feature vector (CFV)** is introduced to address the goal of complete description of watermarks. The complexity of watermarks is demonstrated with several examples. The structure, composition and use of the CFV are demonstrated with ability to speed up searches and retrievals of watermark database.

**Chapter 8: Summary** presents an overview discussion of the work presented in this thesis. Conclusions for capabilities of the proposed original framework for automated identification of watermarks based on chain code are presented. The achieved results are used to illustrate the potential of the new method and to underline the opportunity for possible improvements. Future directions regarding this area of research are listed leaving the proposed framework open for improvements and optimization. The hope is that this method will become a useful tool in filigranology.

## CHAPTER 2

### PAPER AND PAPER WATERMARKS

#### **2.1 The history of paper**

Paper was what irreversibly converted societies from “the oral to the written word”. “Paper – Paging through history”, by Mark Kurlansky [11], gives a well-justified foundation of understanding the importance of paper-making as one of the major technological advances of humanity. The book guides the reader through centuries of inseparable co-development of human societies and their abilities to make and use paper. As the author puts it “Studying the history of paper exposes a number of historical misconceptions, the most important of which is this technological fallacy: the idea that technology changes society. It is exactly the reverse. Society develops technology to address the changes that are taking place within it.” [11, Prologue, pp. XIV]

Many other books make significant contributions to the body of knowledge about paper and paper-making [1, 13, 14, and 16]. Because of the quality and significance of the work of Dard Hunter, his book of “Papermaking: The History and Technique of an Ancient Craft” [1] is a leading example in many reference lists. Hunter’s famous book was published in 1945 and numerous editions continue to be valuable sources with very detailed information. In this thesis there will be multiple referrals to different sections of Hunter’s book. In his first chapter, “Before Paper”, Hunter points out, as Chinese have claimed for centuries, that paper was invented in the year 105 CE by an eunuch named Ts’ai Lun [1, 32]. This reference is quite common. However, the truth has been challenged (and the legend ruined) by twentieth-century archaeological expeditions’ discoveries of paper materials in the dry deserts of Central Asia which predate the CE.

The exact discovery date of paper as a new surface for writing is not of specific importance for this thesis. We may use it as a reference to point that ever since the time of its invention, paper coexisted with other materials for writing: papyrus, silk, tapa, parchment, vellum and others. The list of essential reference books on the subject of papermaking and the history of paper is given in the Bibliography of this thesis [1, 11, 13-16]. Many additional sources are referenced throughout the following text.

From China, papermaking finds its way to Korea and the Japanese archipelago in the sixth century. In the seventh century, paper became known in Vietnam and India. Papermaking was a well-guarded secret and a source of lucrative income for the Chinese empire. A turning point was the battle of Talas in 751 when some Chinese papermakers were captured and possibly saved their lives by revealing their craft to the Abbaside Caliphate [17].

The Arabs learned how to make paper which they called *kaghid* or *qirtas*, words believed to be of Chinese origin. Through Syria, Egypt and the Northern coast of Africa, the craft of papermaking reached Morocco. By the eleventh century, the city of Fez in Morocco had become an important paper producer. The first mention of papermaking in Andalusia, today's Spain, is in 1056 in the town of Xativa. Historically, both Spain and Italy claim to be the first European manufacturers of paper [33]. Crossing the Atlantic in 1620, the English Separatists (Pilgrims) on the Mayflower did not bring papermakers or any printing materials. Paper was not made in America until 1728, when the Liberty Paper Mill became operational on Neponset River near Boston. One curious fact, at least 26 copies of the original Declaration of Independence are found and some of the paper used was not made in America [13, Ch. 11].

In Chapter XVII of his book [1], Hunter provides a very detailed "Chronology of Papermaking, Paper, and the Use of Paper". He starts the chapter by giving due respect to Joel

Munsell, an American printer and publisher for his pioneering work in the history of paper and contemporary recordings in the middle of the nineteenth century. Starting from 2700 BCE, the chronology goes to 1945 when the first edition of the book was published and “...it is stated that there are more than 14,000 paper products.” [pp.584]

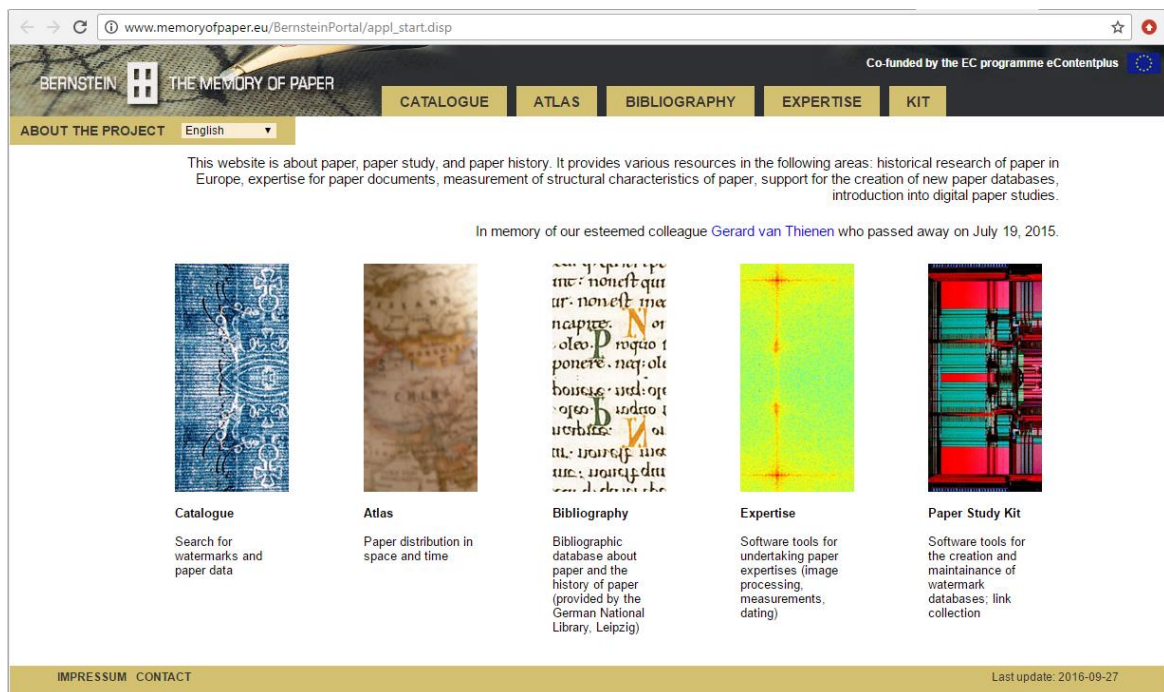
Many scholars studied the history of paper and paper making. The Bernstein Consortium [18] was formed to bring together all major European actors in the field of digital historical paper expertise coming from both humanities and computer sciences (Figure 2.1).



**Figure 2.1: Partners in the Bernstein Consortium** ([www.memoryofpaper.eu](http://www.memoryofpaper.eu))

The project consists of nine partners from six countries, among which the largest collections of paper and watermarks. Other institutions which have contributed prominently to

the project are: Istituto Centrale per il Restauro e la Conservazione del Patrimonio Archivistico e Librario, Rome, Italy; Paper and Watermark Museum of Fabriano, Fabriano, Italy; Institut Valencià de Conservació i Restauració de Béns Culturals, València, Spain; and The State Historical Museum of Russia, Moscow, Russia. As stated in the mission of the Bernstein project, the goal was the creation of a European integrated digital environment for paper history and expertise. Now, the Bernstein project connects all European watermark databases accessible through the Internet. It offers a comprehensive and unrivalled information source about paper. An additional substantial goal was the dissemination of the achieved results to a broad audience in the form of a series of exhibitions, a book about paper history and watermarks and an easily installable software package for paper cataloguing.



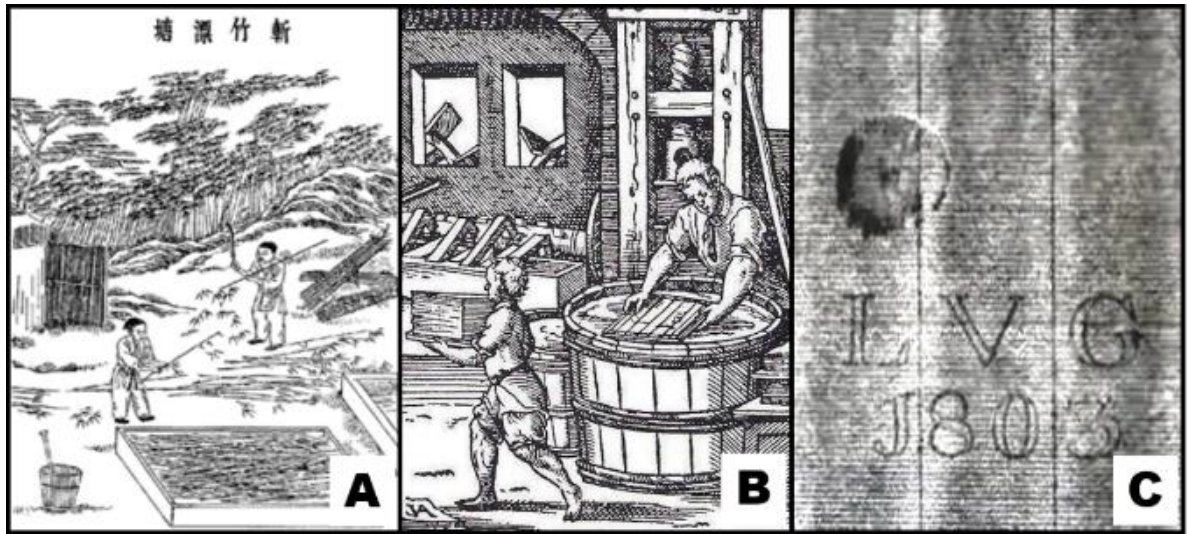
**Figure 2.2: The Memory of Paper integrated environment ([http://www.memoryofpaper.eu/BernsteinPortal/appl\\_start DISP](http://www.memoryofpaper.eu/BernsteinPortal/appl_start DISP), last accessed Sep. 2016)**

The integrated workspace [18] of the Bernstein project provides the digital environment necessary for the integration of resources. It is an Internet application interfaced in six languages (English, French, German, Italian, Russian, and Spanish) that gives access to all Bernstein resources, of which the main components are the catalogue, the atlas, the bibliography, the expertise, and the toolkit (Figure 2.2). Recently, the watermarks in the Bernstein project was sorted also in Hungarian and some in Portuguese, in addition to the six languages defined in the IPH Standard: English, French, German, Italian, Russian, and Spanish.

“Memory of paper” is a deserving name for a technology achievement that helped to preserve so many memories of humanity. Being an excellent material for documents and books, paper cannot survive forever. Digitizing the paper content will insure its preservation for the future. In the Bibliography of his book, Hunter includes “both historical and technical material” about the paper and paper making. He references “Two hundred works on papermaking” of the Oriental, Occidental, Decorated papers, and the Watermarks, which are the topic of the next section.

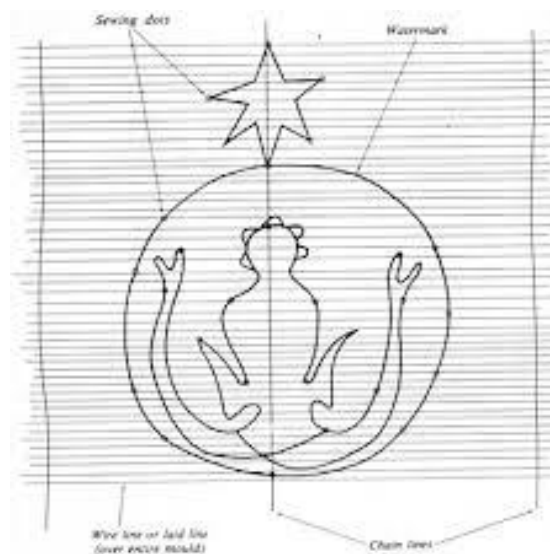
## **2.2 The origin and evolution of paper watermarks**

Due to the changes of the original processing method used in China (Figure 2A), most historical documents and manuscripts made in Europe (Figure 2B) inherently contain *laid* (horizontal) and *chain* (vertical) lines as an underlying artefact (Figure 2C) from the papermaking mould (also often spelled mold). The details of papermaking and the historical progress are well described in [1], [13] and [16].



**Figure 2.3: Early production of paper in China during Han dynasty [12] (A) and Europe [24] (B). The chain and laid lines are visible interference in the watermark image [16] (C).**

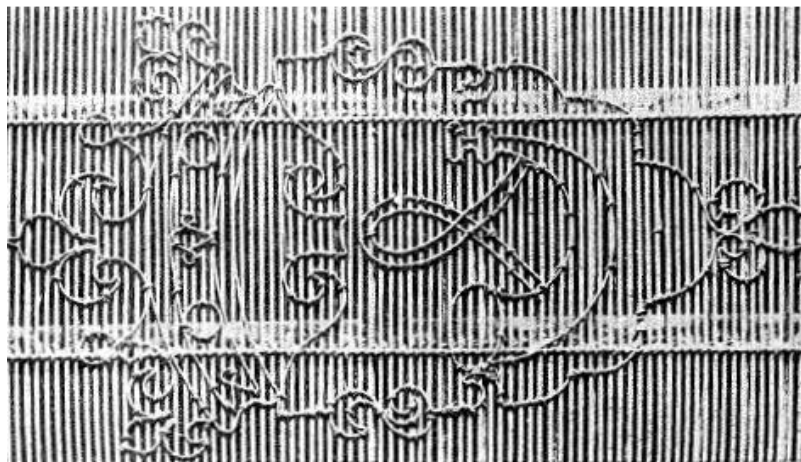
Around 1755 moulds with finely woven wire mesh were introduced and the paper made with them as known as wove paper. Soon after the paper production began in Europe, some manufacturers started to form simple shapes made with wires added to the mould. The additional wires leave an impression with a desired pattern on the finished paper product. The resulting effect was an embedded “wire-mark” left during the pulp’s drying process and historically have become known as “watermark” (Figure 2.4).



**Figure 2.4: An example of a wire figure added to the chain and laid mould wires.**

Only the moulds with metal screens affixed to a wooden frame leave the *laid* and *chain* marks on the paper. A different mould design with cloth screens affixed to a wooden frame continue to be in use even today in Asia. This type of mould is floated in water, the pulp is poured onto the partially-submerged screen and distributed evenly across the surface. The mould is removed from the water and the paper dries directly on the cloth without any marks on the final product. In China, Korea and Japan, a flexible screen is often used with the support of a rigid frame. The screen, which is usually made of bamboo, can be removed from the frame to transfer the formed pulp sheet to a different surface for drying. The level of watermark embedding depends on the quality of the mould. The Japanese *nagashizuki* papermaking uses very fine screens and sometimes a thin cloth, to produce especially fine papers.

In general, the papermaking moulds in Europe use wires, which are closely aligned or woven into a mesh. Through the centuries the mould designs have been modified with the goal to produce better quality paper. It is possible that the wire mould was an Arab idea. However, the papermakers of Fabriano, Italy, are the ones who perfected it. Figure 2.5 displays an image of a hand-mould showing the *verge* and *warp* wires used to secure the complex watermark figure to the *chain* and *laid* wires.



**Figure 2.5: Hand mould with relatively complex watermark. The verge and warp wire are visible attaching the wire figure to the chain and laid wires [16, pp221].**



It is possible that the cause for the first watermark was an accidentally bent mould wire. The artefact was noticeable when the paper was held up to the light. The papermakers realized that now they have the means to “sign” their work. Originally, watermarks may have served to identify papers produced by different workmen within a factory to mark specifically their production for which they were “paid by piece” [2]. So the watermark belonged to the individual papermaker and not the paper mill. As the paper manufacturing process spread across Europe, watermarks were used to indicate the paper brand quality, format, strength, and guidance for folding and cutting of the large sheets produced at that time. The watermark patterns and composition evolved as a reflection of the commercial and cultural exchanges between the middle-age cities [3, 4].

Hunter has discussed some of the existing theories for the usage of watermarks in the early days [1]. In addition to the ones that were already presented here, Hunter comments their use as an identification mark for the sizes of moulds used for papermaking. Or that they were possibly used as symbols of religious groups called *Albigenses*, who used watermarks to identify the members of their group. Another theory suggested that these watermarks were introduced as an expression of artistic skills. A further theory for making watermarks was to help workmen to identify the moulds. The historical records show that Louis de Tignonville, bailiff of Troyes and Charles IV (1398-1409), made the use of watermarks mandatory [1].

The physical creation of watermarks is described in detail by Eakins et al. [19] as follows:

*“A watermark is formed by the attachment of a wire design to the mesh surface of the papermakers’ mould. During paper production the paper pulp is scooped from a vat onto the surface of the mould and the excess water is allowed to drain away through the mesh. The wire watermark, which sits proud of the mesh, reduces the density of fibers deposited on that area*

*of the mould, and when the finished sheet is viewed with transmitted light, the area where the wire had been present is thinner and appears lighter than the remainder of the sheet.”*

In his book [1], Hunter has dedicated three chapters to watermarks: IX, X, and XIV. He described the “Ancient watermarks: six and a half centuries of mystic symbols”, “Later-day watermarks: the nineteenth-century development of watermarks into an artistic and technical achievement” and “The watermarking of machine-made papers and the use of watermarks in detecting forgery”. In these chapters, Hunter has discussed the classification of watermarks from early days until the eighteenth century in four classes, based on their shapes. The first class includes the early watermarks, which have the forms of crosses, ovals, circles, knots, triangles, etc. The second class consists of shapes of the human figure, including a whole body and human parts, such as head, feet, and hands. The third class consists of flowers, trees, leaves, vegetables, grain, plants, and fruits. Finally, the fourth class includes wild and legendary animals, such as unicorns and dragons, as well as snakes, fish, snails, turtles, crabs, scorpions, and varieties of insects. This class also includes bulls’ heads, dogs, camels, elephants, leopards, goats, lambs, cats, horses, deer, and a large variety of birds. Examples of animal watermarks, with type, date used and description are in [20]. Further types of watermarks are given in the site of the International Association of Papers Historians [22]. Watermarks and their relation to the papermaking process is described by Loeber in [7].

Many online databases exist with hand drawn watermarks at known scale and accompanied by textual description of its creation date, origin, manufacturer, etc. [5, 6, and 23]. Not surprisingly, some paper-makers started to copy popular watermarks from their original owners. This led to the introduction of the “countermark” [7], located opposite to the main watermark, usually smaller and representing the initials of the papermaker’s name. An example

would be similar to the LVG initials in Figure 2.3C. It was the use of watermarks in bank notes and stamps that first inspired the use of the term “digital watermarking”.

Modern technology makes the digitization and image storage of historic documents easier, better, and more accessible. The Internet becomes an irreplaceable tool for accessibility of databases, exchange of ideas and cooperation. Many libraries, universities and research centers have developed and continue to expand the use of modern technology to improve the method for recording, processing, and storage/retrieval of digitized documents in general and watermarks content in particular. Two examples are The International Society of Paper Historians portal (IPH) [22] and The Bernstein – Memory of the Paper project [18]. The Bernstein web site has links to more than 30 watermark databases, including:

- WILK - Watermarks in Incunabula printed in the Low Countries, created by Koninklijke Bibliotheek;
- Piccard Online - The entire “Piccard” repertory of watermark tracings, created by Archive of the State of Baden-Württemberg;
- WZMA - Watermarks of the Middle Ages, from Austrian manuscripts of the late Middle Ages;
- Database of Watermarks and Papers Used for Prints and Drawings, created by Nederlands Interuniversitair Kunsthistorisch Instituut (NIKI).

In addition to the above links, The IPH portal has links to:

- Watermark Database of the Dutch University Institute for Art History;
- A Digital Catalogue of Watermarks and Type Ornaments Used by William Stansby in the printing of The Works of Benjamin Jonson;
- The WWW Watermark Archive Initiative;

- Imaging and Watermarks of Rare and Fragile Books, Stamps.

The paper watermarks and their description are the objects of study presented in this thesis. More specifically, paper watermarks found in historical documents and manuscripts. In comparison, the term “digital watermark” is the term to indicate methods for marking, typically in an obscured way, the content of a document. Digital watermark was first used by Komatsu and Tominaga in 1988 [24]. Since then, the concept of digital watermarking has been a subject of intense research and continued to evolve to identify, authenticate, and protect digital content [25]. Further reading on paper watermarks and digital watermarking can be found in [5, 33, 75, 83, 105, and 148].

### **2.3 The study of watermarks**

Paper watermarks found in historical documents and manuscripts are an important part of human heritage. In 1907, C. Briquet (1839-1918), published *Les filigranes* [8] in Geneva, Paris and Leipzig, introducing a collection of 16,000 watermarks found in paper produced before the seventeenth century. He visited hundreds of paper mills in order to compose his collection [8]. The oldest watermark discovered by Briquet printed on paper was produced in 1292 in Fabriano, Bologna, Italy. As previously shown, catalogued as #5410 watermark (Figure 1.1, left), it is the oldest known water mark. Also shown in Figure 1.1, the watermark with catalog #12021, is one of the oldest known watermarks with the name of the paper-maker [26].

According to Hunter [1], the use of the term *Watermark* in English begins in the eighteenth century. In 1708, the New English Dictionary gave the following explanation for the watermark: “*The name was probably given because the watermark, being less opaque than the rest of the paper, had the appearance of having been produced by the action of water.*” In

German the word *Wasserzeichen* was used in the first part of nineteenth century, *Papiermerken* in Dutch, *Filigrane* in French, and *Filigrana* in Italian and Spanish. The last terms are more reference to the “wire figure” rather than the impression it leaves on the paper. It is also the foundation for the term *Filigranology* (the Latin *filium* and *granum* combined with the Greek λογία, logia) -the Auxiliary Science of watermarks’ History. In this thesis, the term watermark is used as reference to the markings embedded in the paper by the sieve used in the paper production method at a given time. The marks from the sieves and the intentionally weaved additional wires could and are used extensively for time period dating of historical documents and manuscripts, which is further described in this thesis. However, dating documents is only one of the applications.

In “Catalogue of Watermarks in Italian Printed Maps CA 1540 – 1600”, by David Woodward [15], the author points out that “In any work involving watermarks, it is necessary to understand that it is the paper mould, not the paper, and certainly not the map or print, that is being identified by the watermark.” [15, pp. 8]. This being said, paper is the carrier of the watermark, the forensic “sample” used to make the discoveries.

Curl Buhler postulates five hypothetical assumptions for any study that uses watermarks to date early documents [27, Pp. 2-3]:

1. A paper mould had a limited life span ranging in four months to four years.
2. There were no extensive time gaps in the use of the mould.
3. The paper produced by a given mould was used up within two or three years of production.
4. Paper was used in the same order in which it was produced.
5. Stock of paper was not accumulated – it was too expensive to be kept unused.

These assumptions are important in regard to using paper making dating and its use for printing. It is not difficult to agree with assumptions 1, 3, and 5. To accept the second assumption for example, we have to ignore that a paper mills did not run without interruptions. Depending on flowing water for power, the mills stopped during dry seasons and winter times. Insufficient supply of rags (key materials for the production of paper) also was the cause for interruptions. In his studies, Allan Stevenson shows that paper mould goes through identifiable stages of aging and changes [28]. He even demonstrates the rate of movement of the watermarks at about a millimeter per month of use [28, pp. 251]. While assumption 2 is important to possible discrepancies in dating of printed materials in general, assumption 5 is especially relevant to maps. One has to consider that only a ream or two are sufficient for printing of maps while book print-shops use 20 to 25 reams per an average size book (one ream has 500 sheets of paper.) It is quite possible that map printing shops used the paper with the same watermarks for a longer period of time.

Scholars consider that the first stage of the study of watermarks started with the first known collections. A brief summary of the watermark studies indicates the tremendous work competed my many scholars. The references may lead to the conclusion that we are historically ready to enter a new, multimedia reach and computationally enhanced stage.

The famous work of Charles Briquet, *Les Filigranes*, was published in 1907, a period which is considered the third stage of the *Filigramology* [29]. This is the stage during which the study of watermarks became independent scientific discipline with the published works of Aurelio Zonghi (1893) and Ernest Kirchner (1896). Humphrey Wanley (1672-1726), a librarian to the Earl of Oxford and an antiquarian, produced a collection of watermarks from 1000 sheets of old papers. Since 1699, the collection is stored in the British Museum. During

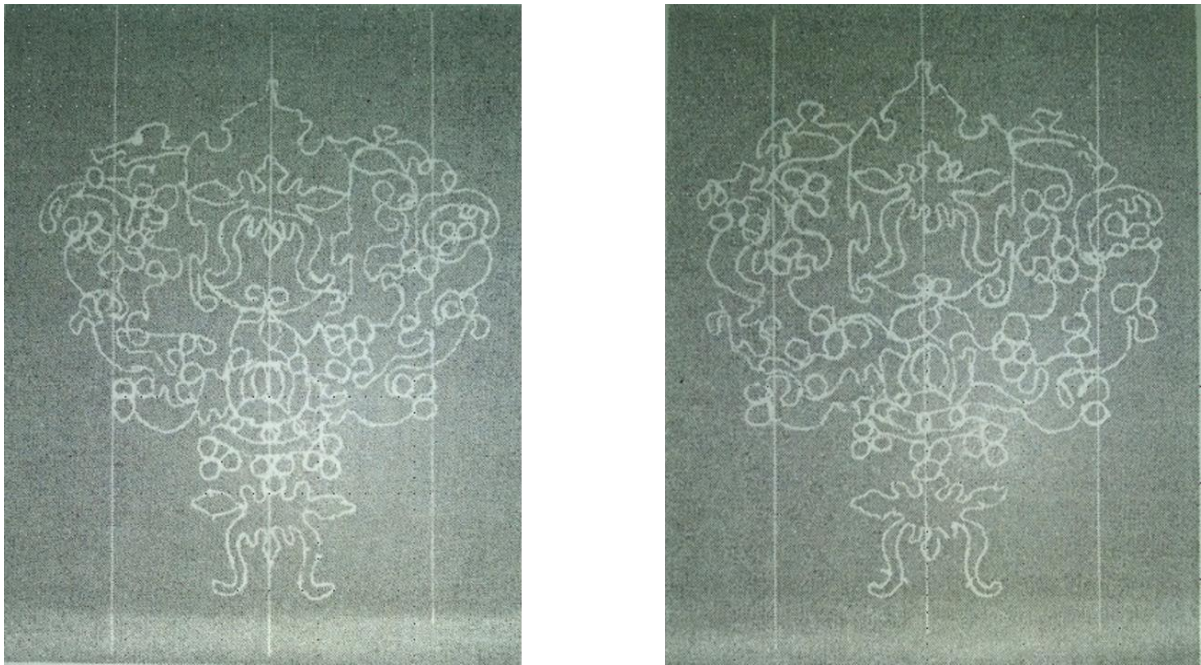
the eighteenth century, the interest towards the watermarks and the collections has grown. Many authors have contributed to the field: Johann Herring (1736), John Levis (1737), Samuel Engel (1741), Joseph Ames (1749), Murray (1763), Rouget (1772), Cristofor Murr (1777), Breilkoph (1784), John Fen (1787), Schwartz (1793), Samuel Denne (1795), Camus and Sardini (1799). The work of collecting watermarks (first stage) continued during the first half of the nineteenth century with the collections made by Gotthelf Fischer (1801), M.C. De la Serra Santander (1803), Jansen (1808), Koning (1816), Hassler (1844) and F. Gutermann (1845).

The second stage in Filigranology has started with the first reproduction of 1493 watermarks in Moscow in 1844 – the Tromonin's watermark album. The same year, the work of S. Sotheby with 265 watermark reproductions was published. Records indicate that scholars from Italy, Germany, Poland, and Romania started to copy watermarks and to publish in specialized journals. The first work about paper mills and their watermarks was completed in Hungary and Transylvania and was published by J. Kemeny in 1844.

The scholars agree that watermarks offer reliable information on how to date paper sources of the late Middle Ages and early modern times. They also agree that while the evidence from watermarks is useful, it needs to be used in combination with additional bibliographical support. Woodward [15] provides an example of using watermarks with offset evidence to reconstruct family of printing plates used in the same place at the same time. In many cases with missing explicit date on the document, watermarks allow close period dating. The process includes finding an identical watermark taken from an already dated source. Because paper was an expensive commodity which was bought and used up within a couple of years, dating based on the identical paper watermarks is a valid scientific approach. One famous example is

the false 1619 year given to the Shakespearean quartos published by Thomas Pavier. Three centuries later in 1908, based on the watermarks, Sir Walter Greg has proved that the quartos were actually published in 1600, 1608, and 1619 [30]. His findings have been confirmed by Allan Stevenson [31].

We have to note that since early seventeenth century paper was often made with pair of moulds with similar but not necessarily identical watermark designs. Double moulds and divided moulds in the eighteenth century are the origin of many twin watermarks [30, 33]. Development of proper imaging systems has facilitated their discovery [32]. Figure 2.6 displays “pair of twin” watermarks, circa 1640, from the Swedish National Archive [16]. Although very similar, these decorative watermarks have distinct differences in their details.



**Figure 2.6: “Pair of twins” watermarks, c. 1640, Swedish National Archive [16].**

Atanasiu [34, 35], as part of his work in the Bernstein project [18], developed two applications which helped in studying laid lines. The first is for laid line density measurement, known as “AD751”, which locates the frequency of these lines after Fourier transform [36],



and the other is for laid lines suppression and extraction, known as “Blue Nile” project. Three other useful applications are: 1) “Filigrana” (another laid lines density measurement tool); 2) “Watermark Scissors” (an application which segments an image which contains a number of watermarks into smaller images according to the number of watermarks); and 3) “WMT” (an application for interactively measuring width and height of watermarks) [35, 36].

It is important to note that during the life span of a mould the same watermark may produce slightly different impressions due to deterioration and shifting of the wire. The result is a variant of the same watermark known as “states” or “variant states” (Figure 2.7). Their differences are considered “false variants” to distinguish them from the “genuine variants” [16].



**Figure 2.7: Different states of the same watermark. Swedish National archive, c 1825. The “F” has not changed while the “C” has shifted and changed shape noticeably [16].**

In addition to the watermarks as artefacts of the manual papermaking, watermarks also exist in machine-made paper. The International Association of Paper Historians [22] has created a taxonomy of terms for describing the components of paper, including the paper watermark. Each watermark is assigned a code (e.g. E8 for snake), and these codes are arranged in organizational tree structures, (e.g. Birds → Eagle → Double-headed).

The First International Conference on the History, Function and Study of Watermarks, has discussed the importance of watermarks and respectively their study. The proceedings of the

conference were published in [37]. In addition to *Les filigranes* by Briquet, there are several others published catalogues of watermarks. Collections can be found in [38 - 42]. A list of books of reproduced watermarks by tracing is in [81], together with a number of traced watermarks in each book. An interesting and rare find is the reference to Soviet era watermark studies [43].

Multiple web-based archives of watermarks are easy to access and use. Some examples include [44 - 52] and an extensive list of links to web databases is compiled in [18]. It is important to note again that easy access at any time is one advantage of digitized documents in addition to preserving digitally the watermarks from their paper carrier decay. The digital format allows many researchers to design and develop new methods for watermark processing, enhancement, and description.

Going back to the facts about the original Declaration of Independence, the paper used to print it was (and still is) Dutch with three different watermarks. Some of the paper has the L. V. Gerrevinck watermark (Figure 2.3C, shown previously). It is the same watermark found on the notes of Tomas Jefferson for the 1776 Virginia Constitution.

### **2.3.1 Imaging and reproduction techniques for watermarks**

Dating of historical documents is only one compelling reason for the continuing research on watermarks. Beautiful in their simplicity and enigmatic in their eternity, watermarks may stay hidden for centuries waiting to be discovered and reveal valuable information [28, 37]. It is very useful to humanity to preserve the cultural heritage of watermarks as stamps of history and to make them accessible to everyone interested without the physical access to the original paper resource. The problem of converting the watermark visible image from the paper into a

meaningful piece of information is still present in the study of these valuable cultural artefacts. Among many organization dedicated to the understanding and preservation of paper watermarks, The Centre for the Study of the Renaissance at the University of Warwick, Coventry, UK, has many useful links to libraries and online watermarks databases [5]. *The Memory of Paper* website of Bernstein, has a map of the centers with digitized and not digitized documents with watermarks (551,138 records, 36 databases) [6].

There are many ways to gather identifying information about a watermark. Being visual artefacts and not-linguistic data, imaging watermarks is a better way of retaining information compared to textual description. Modern imaging techniques have facilitated the digital image acquisition of watermarks and the number of digitized watermark collections continues to grow. Nevertheless, the estimate in [6] is that about 66% of documents with watermarks are still not digitized. Attempts at standardizing the subjective description of watermarks has been published and is also available online [21, 22]. The watermarks are classified according to their thematic topic (e.g. flower, lamb, eagle, mill, cross, and others) in multiple languages. Due to their high numbers, variety or rarity, and not the least to the subjective nature of existing classifications, matching of a newly obtained (and perhaps unknown) watermark is a difficult task. Interoperability of existing databases is a separate and very challenging problem as well.

Modern technologies facilitate imaging, processing, and storage in digital format of many cultural treasures and their preservation for the future generations. Captured digital images can be compared, processed, stored and retrieved easily. Numerous algorithms allow further image enhancements to be easily applied. Modern cameras provide captured images of high resolution, which allows the observer to see very small details. Various projects, such as AHDS [53], Minerva [54], Pulman [55], and MUSICNETWORK [56], provide guidelines for some

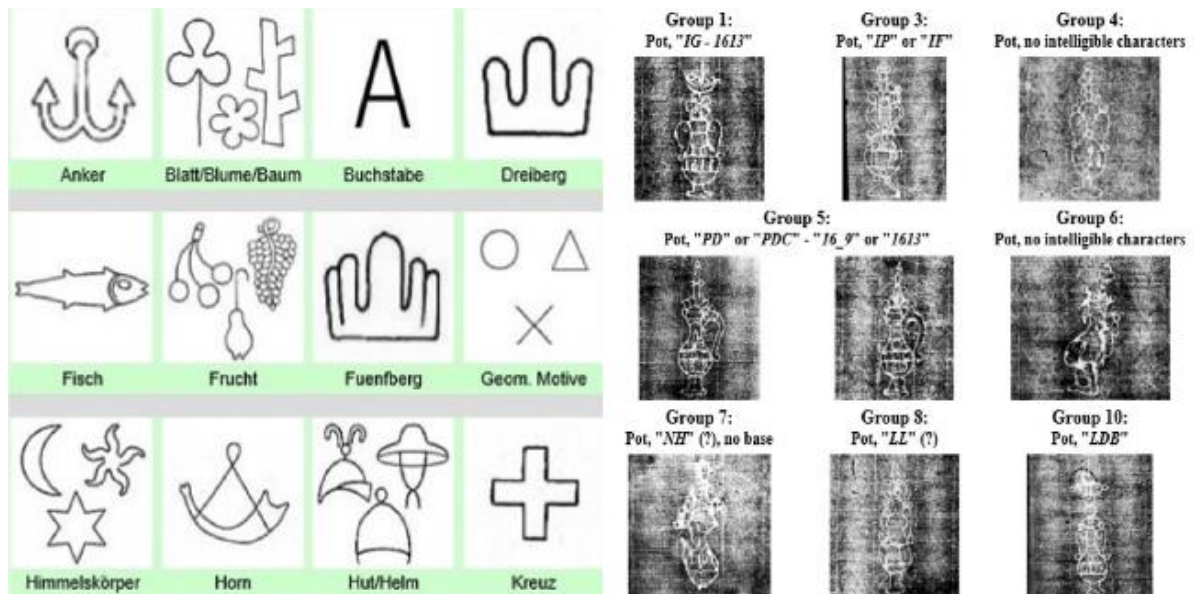
of the most useful digitization techniques. Existing methods for paper watermarks reproduction include: manual tracing, rubbing, Ilkley, phosphorescence watermark imaging, Dylux 503-1B503 photosensitive paper, transmitted light photography (backlight imaging), thermography, Beta-, electron-, and soft X-ray radiography imaging [57 - 65]. The SHREW project (“SHape REtrieval system for Water-marking images”) compares different watermark digitization techniques [66, 67].

The manual tracing of watermarks takes time, needs certain skills, and does not capture some watermark details. In addition to the subjective component, it is an intensively contact method which is a concern especially for original documents in fragile state. Dylux and back-lighting methods provide good quality results with watermark details. However, the reproductions contain undesirable interferences (e.g. laid and chain lines, paper stains and defects, additional writings and others). Using Dylux is associated with “off-gassing” from the chemicals used in its production [57]. Backlight imaging followed by various image processing techniques is one of the most popular approach for watermarks digitization. The method provides accurate results, with optimal efforts in respect to cost and time. Main problem of using backlight imaging is the presence of the front-side (*recto*) and back-side (*verso*) prints in the acquired image. X-ray imaging and Beta-radiography method provide good quality images, but they are quite complex and require expensive resources and safety measures working with ionizing and radioactive sources [62 - 65]. Modern imaging techniques favor non-contact imaging including hyperspectral imaging with ultraviolet and near infra-red light.

As described in [68], Christie-Miller collaboration with Solar Imaging Systems Ltd. [69] has developed a hardware back-lighting digitization system - Advanced Paper Imaging System. The goal was to record the paper structure (respectively the watermarks) in order to

provide digital identification for stolen manuscripts. Another purpose was to capture and preserve in digital format valuable artefacts. As a side note, the system was designed to allow imaging of bound manuscripts opened at 45°, so the digitization is less intruding and does not damage the manuscripts.

After digitization and segmentation the watermarks images are with different quality, size, resolution, and orientation. Figure 2.8 displays watermarks from two different collection. On the left, is a sample from the searchable online watermark collections [69]. On the right, are watermarks from Primary paper by William Stansby in the Jonson Folio [70]. One can see the extreme to which the quality of watermarks representation varies based on the acquisition technique used.



**Figure 2.8: Examples of watermarks digitized with different techniques and respectively different quality.**

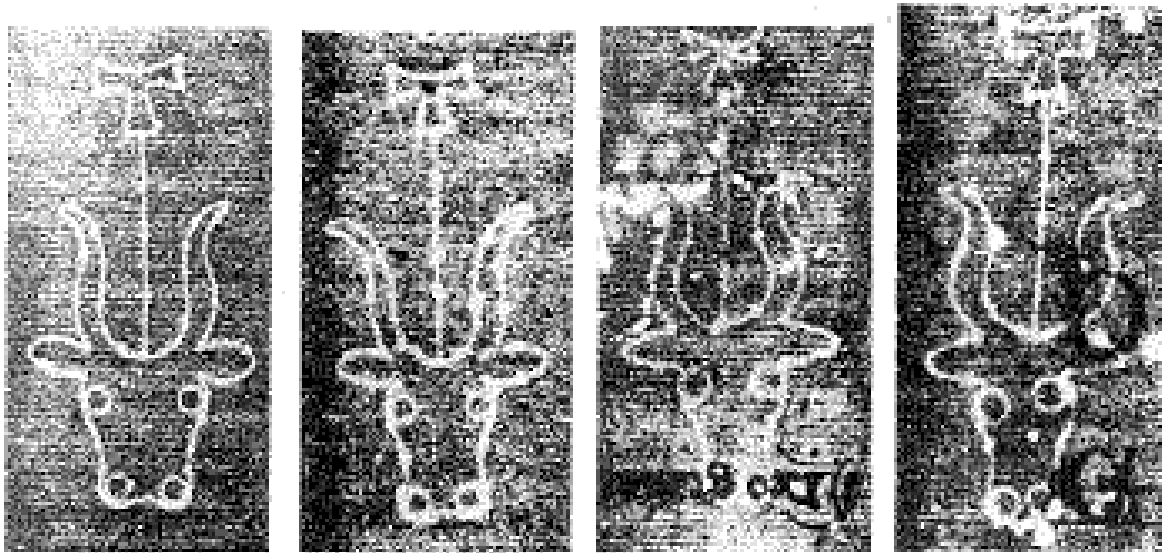
Digital images of watermarks from many sources require extensive image enhancement before filing, classification, and storage. A successful automated search, match, and retrieval would require even more extensive preprocessing before it could be successfully applied. The

quality of the images depend on the document's condition and the imaging technology used. Imaging systems continue to improve (optics, filters, sensors, light sources). The computational capacity of CPUs and GPUs are also improving. There are multiple algorithms developed for image processing and enhancement and many more continue to be developed. However, it is important to notice that while some methods achieve remarkable results, there are no universal approaches that can work in all possible scenarios.

### **2.3.2 Description, classification and storage**

The International Society of Paper Historians have developed a set of criteria that can be used to describe paper with considerable precision [22]. Known as the *IPH Standard*, its section 3.1 includes a series of entries for watermark data that allow for the identification of watermarks based on a set of unambiguous individualizing characteristics: type, dimensions, position in respect to the *chain-lines*, location, and others. The exhaustive set of descriptors specified in the *IPH Standard* are followed by leading scholars in the field of literary manuscripts. It is also well accepted that even the best textual descriptions are never as satisfactory as a high quality image reproduction such as a beta radiograph [65]. There are a number of printed volumes or digital resources that contain reproductions of thousands of watermarks. The available databases are a very important resource and the result of extensive scholarship. But they are not perfect because the most common (and least satisfactory) way to identify a watermark is to cite the known closest match in a database of watermarks based on textual description. This is especially valid for watermarks with similarity, being from the same “family” of watermarks or “the same” watermark from different moulds.

An important ability is to distinguish “identical” from “twin” watermarks. “Twin” watermarks are the result of common papermaking using a pairs of moulds to accelerate the process. The pair of moulds had similar but not necessarily identical watermark designs. The identification of “twins” and their distinction is important for the studying historical documents [33, 71]. Figure 2.9 displays four different “bull’s-head” watermarks reproduced from A. Stevenson, “The problem of the Missale Speciale” [28].



**Figure 2.9: Bull’s head twin pairs, Missale Speciale, c. 1473. Two pairs from left to right: Wide face - Square face and Level eye - Slant eye. Adapted from [28].**

Some of the techniques for description, classification and storage become clearer by reviewing the methods currently used for cataloging, searching and retrieval of watermarks. It is a mutual, two-way relationship in which a better and complete description provides easier cataloging, and searching. On the other hand, the needs for an easier and faster retrieval capabilities, demands better methods for description and classification of the variety of watermarks. The next section provides a review of current techniques for cataloging, searching, and comparison of watermarks.

### **2.3.3 Current methods for cataloging, search and retrieval**

Unfortunately, many different methods exist for cataloging of watermarks regardless the efforts to standardize the approach in order to achieve interoperability. As previously noted, section 3.1 of the *IPH Standard* [22] of The International Society of Paper Historians includes a series of entries for watermark data that allow for the identification of watermarks based on a set of unambiguous individualizing characteristics: type, dimensions, position in respect to the *chain-lines*, location, and others. The exhaustive set of descriptors prescribed in the *IPH Standard* are followed by leading scholars in the field of literary manuscripts. Nevertheless, it is also well accepted that even the best textual description are never as satisfactory as a high quality image reproductions of the watermarks.

The following Figure 2.10 presents the IPH specifications for describing a simple and more complex watermark (IPH Standard, version 2.1.1).



A EXAMPLE OF A SIMPLE WATERMARK

Description in full (according to 3.1.5):

*below a crown, capital italics CLK*

Completely coded description (according to 3.1.4 and appendix I):

R3/I - {b: (i: X „CLK“ )}  
CROWN AND BELOW ITALICS ABBREV. CLK

B EXAMPLE OF A COMPLEX WATERMARK

Description in full (according to 3.1.5):

*Eagle, crowned, overlaid with italic capitals FWR, holding sceptre and sword in the left claw, an orb in the right claw, at the top the name „G W Loeschge“, at the bottom „in Ansbach“*

Completely coded description (according to 3.1.4 and appendix I):

D5/I [ R3/I - {c: (i: X „FWR“) -  
EAGLE WITH CROWN AND OVERLAID WITH ITALICS ABBREV. FWR AND




CL: p06 [ R7 - M14 - {r: p06 [  
LEFT CLAW WITH SCEPTRE AND SWORD AND RIGHT CLAW WITH

R4 - {t: → Y „G W L0ESCHGE“ - {b:  
ORB AND TOP HORIZONTAL NAME G W LOESCHGE AND BOTTOM

→ Y „IN ANSBACH“ } ]  
HORIZONTAL NAME IN ANSBACH

**Figure 2.10: Examples for description of simple and complex watermarks.  
IPH Standard, Version 2.1.1, 2013.**

Figure 2.11 shows an example for cataloging class Lion watermarks with different subclasses. All descriptors labels are given in six languages. IPH Standard describes in details many aspects for the registration of papers with or without watermarks. This includes abbreviation descriptors, dating, origin, and many others.

C17		Lion Löwe Lion León Leone Лев	
	C17/1	Lion (general) Löwe (allgemein) Lion (en général) León (general) Leone (generico) Лев (вообще)	
	C17/2	Lion, Dutch (holding arrows or sword) Löwe von Geldern (mit Pfeilbündel oder Schwert) Lion des Pays-Bas (avec flèches ou épée) León de Paisas Bajos (con flechas o espada) Leone olandese (con spade o frecce) Лев, голландский (держащий стрелы или меч)	
	C17/3	Lion of St. Mark Markuslöwe Lion de St. Marc León de San Marcos Leone di San Marco Лев Святого Марка	

**Figure 2.11: Lion class with subclasses. IPH Standard, Version 2.1.1, 2013.**

In “Catalogue of Watermark” by David Woodward [15, pp. 13], the “Explanation of Entries” for the watermarks are as follows:

- A) *A general description of the watermark mould, including*
1. *The watermark mould number*
  2. *A descriptive title, organized in the order recommended by the International Standard for the Registration of Watermarks published by the International Association of Paper Historians, Provisional Edition, 1992.*
  3. *The size of the salient feature, such as “circle 52 mm.”*
  4. *A notation of special characteristics.*
  5. *A code such as “A1{p05}-U1” following the specifications in the International Standard for the Registration of Watermarks published by the International Association of Paper Historians, Provisional Edition, 1992.*

B) *Each subentry reflects to a specific impression of a map or print, with the following information:*

1. *Date printed on impression; n.d. signifies that no date is printed.*
2. *The title or area of the map {in square brackets when summarized in English}.*
3. *Place of publication printed on impression; n. pl. signifies that no place is printed on the map; [place in square brackets is estimated from other evidence].*
4. *Name of printer or publisher appearing on impression; n. pr. signifies that no publisher is printed on the map; {printer in square brackets is estimated from other evidence}.*
5. *Map number from Toley's list of Italian printed maps.*
6. *Other references, when not listed in Toley or to help to identify the impression.*
7. *The institution and call number where the impression is found.*

The presented referenced text illustrates several aspects: 1) the efforts made to capture the complete information; 2) the complexity and manual labor that is involved; 3) the lack of options for an easy way for search and retrieval.

Many attempts have been made to develop a system for classification and retrieval watermark images, both by textual-based and content-based approaches. Briquet [8] has manually classified about 16.000 watermarks in about 80 textual classes. Del Marmol [72] has classified ancient watermarks based on the date of their creation. Pure textual classification systems might be less prone to errors but very tedious to use. Textual descriptions and watermark labels are inconsistent in methodology and format, erroneous or just not detailed enough. Many watermarks are without labels and cannot be used for textual queries.

Advances in imaging technologies have made possible recent attempts to focus on the real content of watermark images. In [73] the authors describe the set-up of a system capable to

add, edit or remove watermarks with the possibility of textual and content-based retrieval. As described in the paper, the authors used a 16-bin large circular histogram computed around the center of gravity of each watermark image. In addition, eight directional filters were applied to each image and used as a feature vector. The algorithms were tested on a watermark database consisting of 120 images, split up into 12 different classes. The system achieved a probability of 86% that the first retrieved image belongs to the same class as the query image.

In [74], the authors used three sets of various global moment features and three sets of component-based features. The component feature set consists of several shape descriptors which are extracted from various image regions. For testing, 806 tracings of watermarks from the Churchill collection [38] are used and 15 images and their ground-truths have been manually selected. The retrieval system was based on the city-block distance measure. The system was evaluated with normalized and averaged precision and recall scores. The authors report a precision of 0.53 and a recall of 0.81 for their best scores with the conclusion that for their database, global features work better than the component-based (local) features extracted from various regions.

In [75], the authors emphasize that “successful research in watermarks (filigranology) is reliant on the availability of relevant and accurate reference material that can be easily accessed.” After that, the authors describe different technologies for recording watermarks and the necessary criteria for proper recording and classification of the watermarks. Image examples are given for watermarks recording and processing. For image searching method, the authors used content based image retrieval (CBIR). First, *ground truth* images are established. For the retrieval process, a query image is chosen and a search is performed by the system and by independent researchers which are asked to select the most similar images from the database

based on visual inspection. The ground truth sets have been selected from an archive of 2000 digital images of watermarks. In the referenced paper, the evaluation is marked as “ongoing” and numerical performance values are not given.

Textual queries are processed very fast with modern computers but the queries have major limitations. As already mentioned, despite the existence of a standard, textual descriptions are subjective (inconsistent) and often are incomplete: some of the patterns appearing in a given watermark may not be described in the text. For these reasons, queries based on global morphological characteristics as well as on the watermark shape are essential.

The watermarks can be described as shapes of their contours contained in their images. The information can be used for accessing the archived images of the watermarks or pictorial documents with them [22, 36, 67, 76-78]. Each watermark is described by a list of morphological characteristics. The spacing between the chain lines and their density (lines per unit length) helps to determine the type of the paper. The overall size of the watermark is defined by the height and width of its bounding box. The number of closed regions and the shapes of individual regions are described by the invariant moments [79].

During the review process of existing methods for watermarks description, I have made useful observations. When multiple watermark regions exist, the centers of gravity of the watermark’s main region are selected to build a new normalized coordinate frame. The center of the largest region is the origin of the new coordinate frame and the two other regions form the two independent coordinate axes. The remaining regions from the watermark are then described in this new frame by means of two scalar coordinates. This morphological description of the regions is invariant to rotation, translation and scaling. The percentage of black pixels present in the watermark bounding box are used to define the watermark extent.

This value allows to assess the complexity of the watermark and to date its period - more ancient watermarks are simpler with and have less black pixels compared to the more recent ones. The location of the watermark with respect to the chain lines is used to compute two values: the distances between the chain lines and the closest pixels of the watermark. At the time of recording and archival, these characteristics are determined and stored into hash tables. At retrieval times, queries can be made according to one or more of these characteristics.

The *SHREW* (“SHape REtrieval of Watermarks”) is a database project for image retrieval of historical watermarked papers. The input data consists of watermarks traced by Churchill [38] and images reproduced by electron-radiography. The enhanced watermark images are stored in a database and for a retrieval, a given watermark can be matched with previously stored watermarks and similar shapes [19]. Traced watermarked images were processed for feature extraction: images are first converted to binary using a constant threshold, then noise is reduced using filters (e.g., mean, median and Gaussian), images are then enhanced using morphological operators. These enhance images were combined with shape retrieval techniques and evaluated [80]. The *SHREW* was further developed and expanded with datasets reproduced using electron-radiography [81]. Chain and laid lines were removed by applying low-pass frequency filtering. The background was approximated by applying a median filter multiple times to produce an image, which is then subtracted from the original image. Some examples of the processing results are given in [19].

The *SHREW* system has been implemented in an SQL database (*The North Umbria Watermarks Archive*) for publication on the web to facilitate retrieval of information on watermarks. The authors admit the need for further work on image processing to improve

quality and on the content based image retrieval to improve the performance of shape matching and retrieval.

An early attempt for developing a “smart” document retrieval system is described by Salton in [82]. Wenger *et al.* [83] have proposed the *INTAS* project: a Distributed Database and Processing System for Watermarks, which is a multinational effort coordinated by the Austrian Academy of Sciences [84]. The aim of the project is to build a database for watermarks existing in West Europe and Russia, which can be accessed widely, and can help scholars to study the captured watermarks and use them to study undated documents. The objectives of the project outline the important needs and associated challenges for:

- Development of a distributed database system of watermarks;
- Methods and technology for producing watermark hardcopies from manuscripts and early printed books;
- Digital processing of watermark images;
- Integrated system for database management and watermark processing;
- Population of the distributed database;
- Exploitation and dissemination of results.

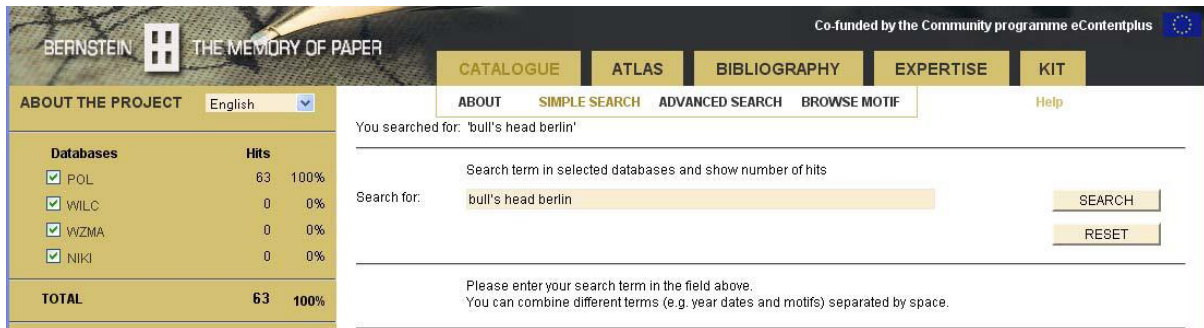
This list can be viewed as a charter for continuous work in the field of digitization of historical documents. The aim of *INTAS* was to improve the reproduction techniques, including radiographic, back-lighting and rubbing techniques, resulted in the birth of the first two electronic watermark databases in Russia [85]. The reproduced images were enhanced (contrast enhancement) and watermark contours were approximated using semi-automatic processes for identification purposes [86]. These enhanced images are then entered into the database. An acknowledgment is needed to emphasize the important role of individual

leadership: the leading author of [83], Mr. Emanuel Wenger, is also the coordinator of the *Bernstein Consortium – The memory of papers* [18]. Similarly to Bernstein project, INTAS aims to create a digital environment for researchers to study paper and to link European databases of reproduced watermarks together, and provide image processing tools to measure paper features. Additional resources can be accessed using the links in [87].

The *Profil* is another watermark database project [88] with the goal to provide scholars the ability to identify watermarked paper. Data was reproduced using beta-radiography in the National French Library. The watermarks were scanned and entered into the database, together with a description of each watermarked document. The recorded images were processed to enhance their quality using different filtering methods (e.g., median, Gaussian filters, filtering the frequency domain).

The Bernstein *Catalogue* allows search in and data retrieval from various online databases. The search can be formulated in the six languages specified in the IPH Standard (English, French, German, Italian, Russian, and Spanish). Recently, queries are possible in Hungarian and some in Portuguese. This facts can be viewed as expansion of the interest in the watermarks as a cultural heritage and as the enabling power of the internet. All search terms are translated automatically into the supported languages according to the Bernstein vocabulary for watermark descriptions. For example, a search for *mermaid* (English) is carried in Italian for *sirena* and also in German for *Meerjungfrau*. The catalogue offers three modes for the search: *simple search*, *advanced search*, and *browse motif* [18]. Examples for each type are given in Figures 2.12, 2.13, and 2.14.

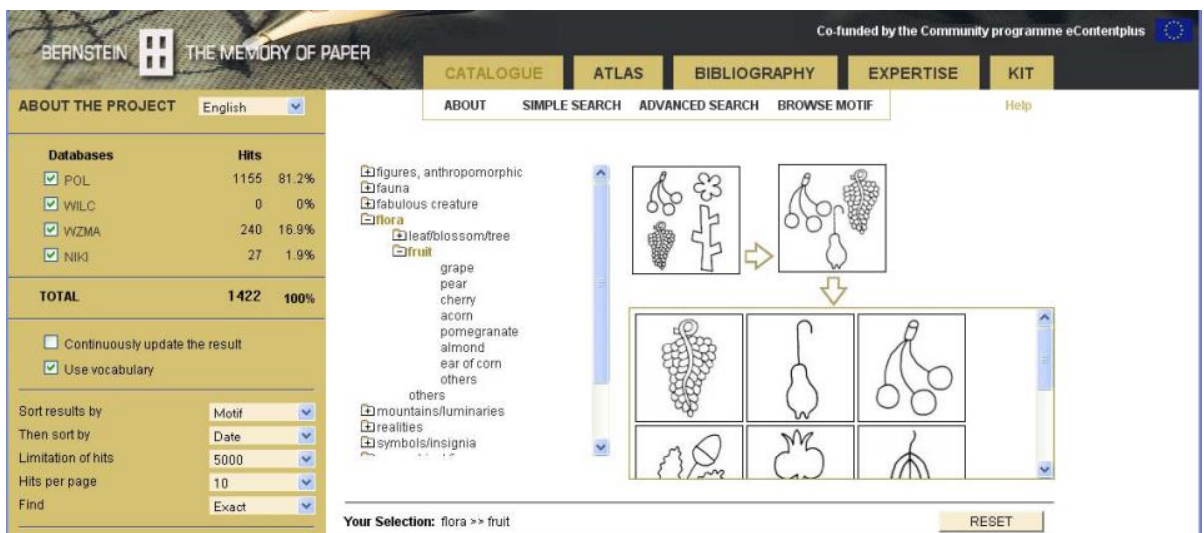




**Figure 2.12:** For a simple search the user just enters the search terms in a search field.

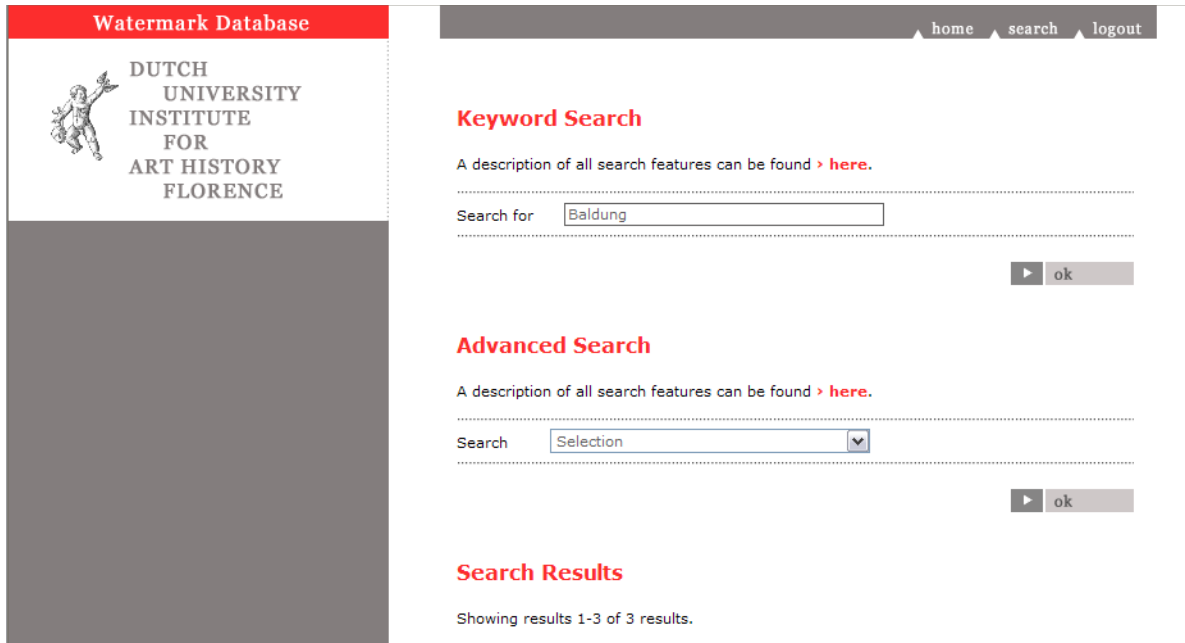


**Figure 2.13:** For an advanced search the user can combine several search fields.

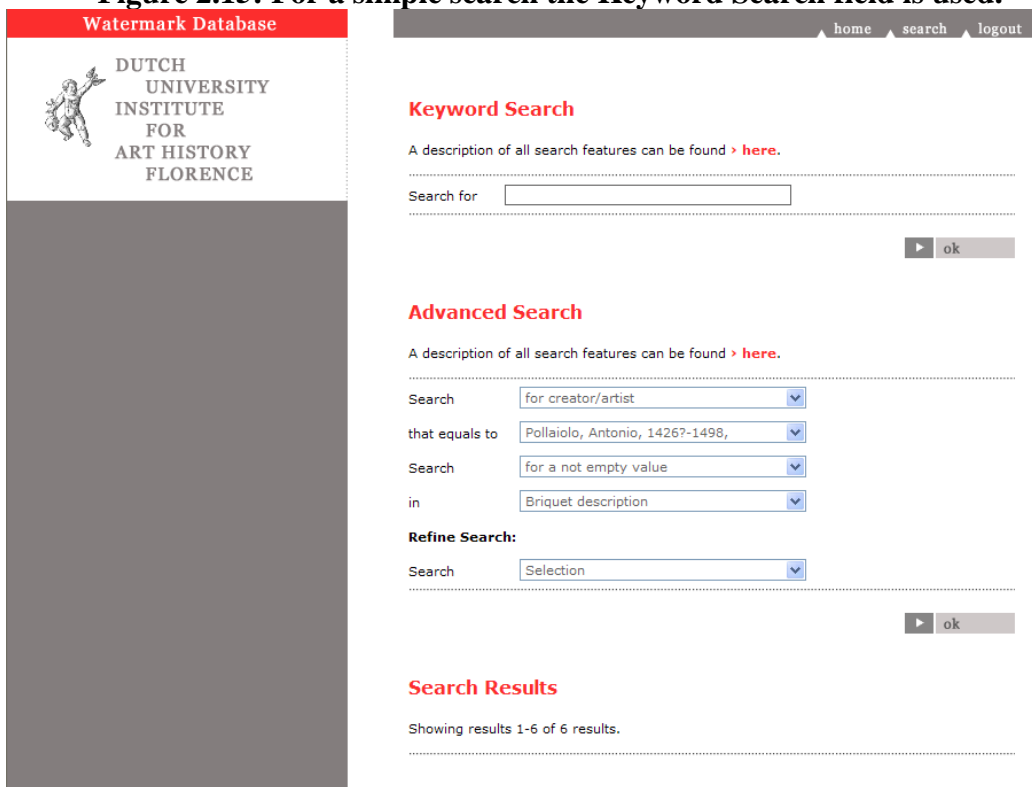


**Figure 2.14:** The browsing with a motif offers the possibility to navigate in the tree structure of the systematics classification by names or by icons.

The Watermark Database created by the Dutch University Institute of Art History in Florence has similar search methods (Figure 2.15 and 2.16).



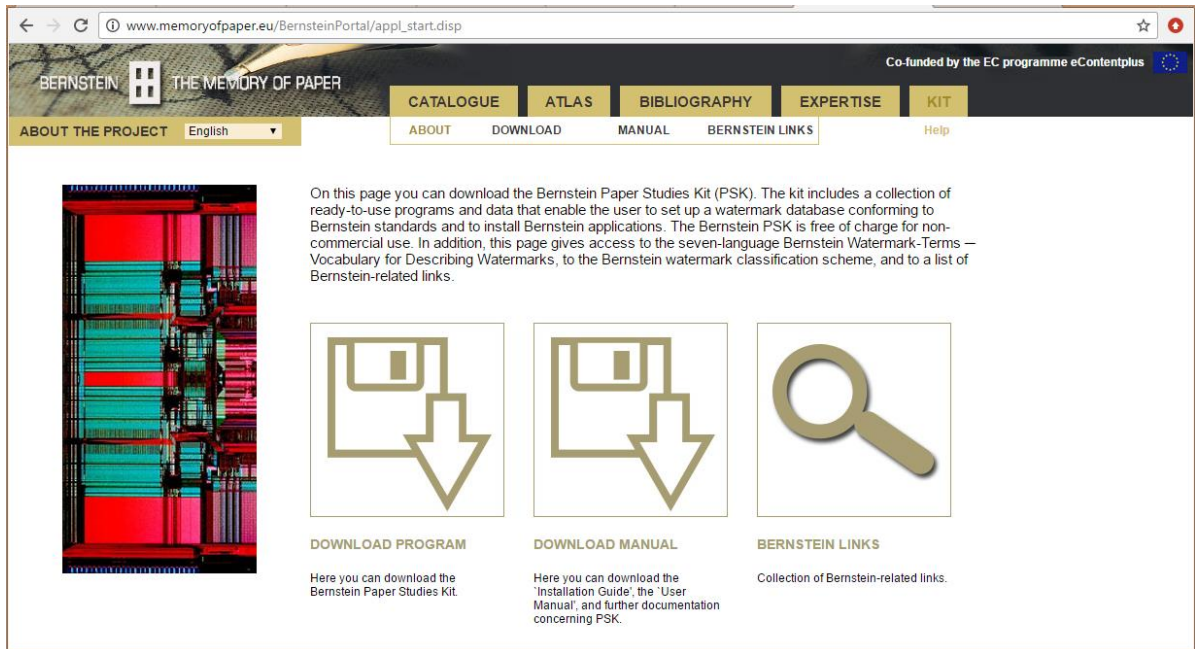
**Figure 2.15: For a simple search the Keyword Search field is used.**



**Figure 2.16: For an advanced search up to 8 parameters may be used: 1. Creator/ Artist; 2. Signatures; 3. Time Period; 4. Keyword; 5. Range; 6. Exact value; 7. Not Empty Value; 8. Empty Value.**

The field of study of historical documents and paper watermarks is wide open and offers many possibilities for further explorations. The internet becomes an irreplaceable tool for

research and cooperation. Almost the entire search for information for this thesis was performed on the Internet. As a conclusion for the options available on line, the reference to the Bernstein Paper Studies Kit (PSK) is presented in Figure 2.17.



**Figure 2.17: Bernstein Paper Studies Kit – links, software and manuals for the studies of watermarks.**

The kit offers a collection of ready-to-use software programs and data that enable the user to set up a watermark database conforming to Bernstein standards and to install Bernstein applications. The PSK is free of charge for non-commercial use. In addition, the portal links give access to:

- the seven-language Bernstein Watermark-Terms (Vocabulary for Describing Watermarks);
- the Bernstein watermark classification scheme;
- Bernstein-related links.

Utilizing the Internet, Rauber *et al.* [73, 89] have proposed a system for the management, archival and retrieval of historical papers in a database accessible on line. The goal is to enable scholars to determine date and origin of unknown paper by being able to compare watermarks of interest with known watermarks in the database. The database contained an image and textual description of each watermark. Back-lighting was used for image capturing, followed by specific image processing algorithms [91] such as contrast and contour enhancement, removal of laid and chain lines and other imperfections from the paper. The work also included scanned images of the watermarks traced by hand by Briquet [8]. The proposed and used textual and image retrieval classifications of watermarks was specified with the following steps:

1. The class of the watermark, as presented by Briquet.
2. The use of IPH code presented by the International Association of Paper Historians.
3. Retrieval by specifying global features, using 12 features (e.g., watermark size, watermark position on paper, spacing between two sequential chain lines).
4. Retrieval by comparing similar images. A similarity task processing algorithm is presented to compare the shape of a given watermark with other watermarks stored in the database: two algorithms were proposed for comparing similarities, Circular histogram and Directional algorithms. Details of these algorithms are in [73].
5. Retrieval by drawing an approximate shape – the authors built the option which allows historians to draw watermarks manually in the absence of an image.
6. Retrieval using small patterns, that is, retrieval using only part of the watermark, where watermarks in the database are indexed into a hash table, and convolution is applied to search for similar watermarks.

In [90], the authors report the continuous work to develop further their system for watermark images retrieval for ancient paper identification with ability for the archival, retrieval, and distribution of electronic documents. The idea is to access information via the World Wide Web with a search mechanism that allows the retrieval of text and images according to their content [91, 92]. The purpose is to design and implement an efficient method for comparing a given watermark with another similar known watermark. In their work, the authors referred previous attempts to develop such systems [26, 93]. The system developed in the project was applied to the watermark collection of the internationally known Swiss Paper Museum in Basel. Developed in the 1980s and early 1990s, the collection of the Museum consists of thousands of images of historical papers as well as ancient watermarks [89].

Presenting their work, the author address multiple important aspects. The watermark's description consists of the textual characterization of the paper containing the watermark, the origin, the date of creation, and a paper is described by approximately 150 different parameters [21, 22]. The reference to the available methods for the process of watermarks identification is a very good illustration of the capabilities and also the limitations of the used methods and the intensity of the efforts for their application. The reported work is one of very few examples when an attempt is made to expand the search, matching, and retrieval methods and include image based feature extraction. For illustration of this pioneering work and to demonstrate some of the constraints, several of the examples for the work are presented here. It is very relevant to the current state of watermark identification and retrieval and creates a good foundation for the framework proposed in this thesis. Figure 2.18 demonstrates retrieval by using a class and Figure 2.19 show the retrieval using the IPH code.



(a)



(b)

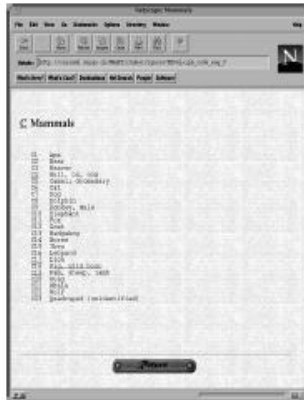


(c)

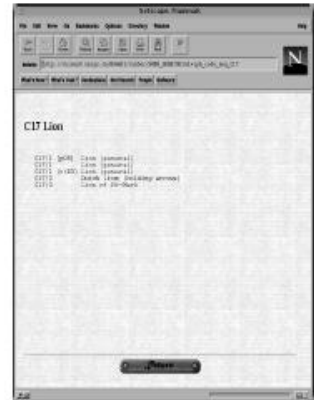
**Figure 2.18: Internet page to access watermarks by using the class as the retrieval mechanism. (a) Textual list of classes. (b) Iconic representation of the classes. (c) Main page for the class Eagle (Aigle in French).**



(a)



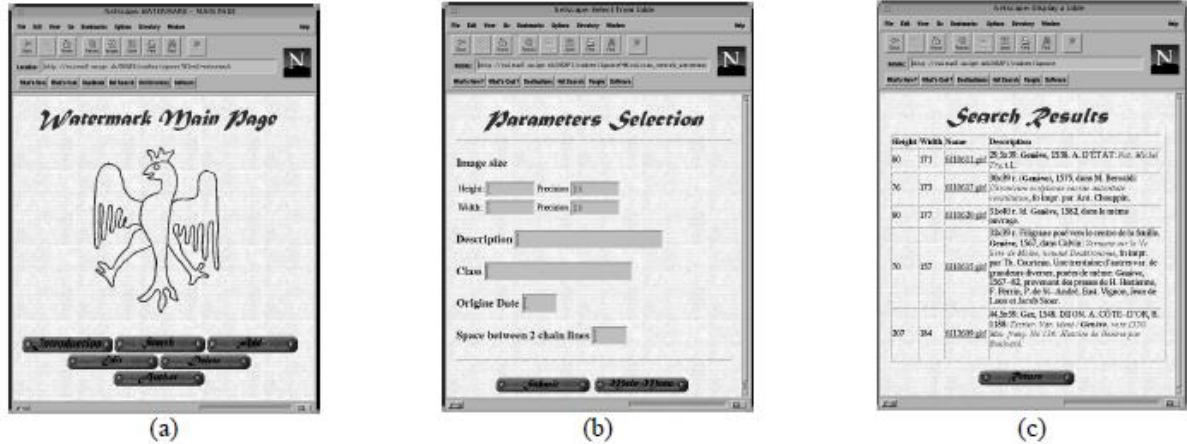
(b)



(c)

**Figure 2.19: Main pages for accessing watermarks by using the IPH code. (a) Main index in French. (b) Sub-class Mammals. (c) Sub-Sub class Lion.**

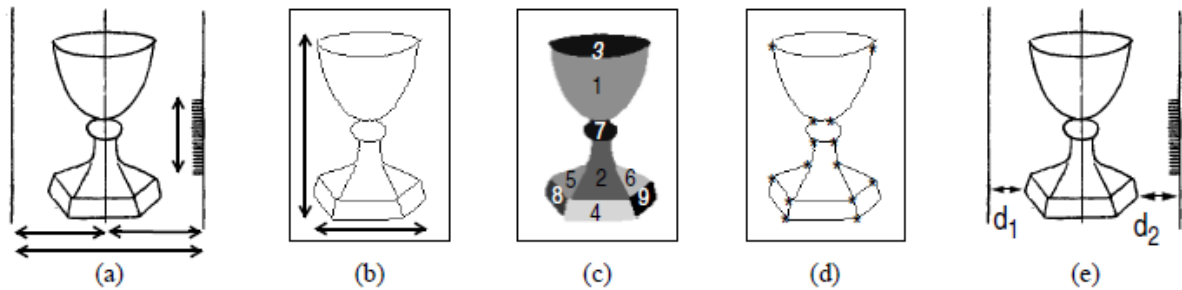
From the main watermark page, the users are prompted to enter values corresponding to the watermark they wish to retrieve (Figure 2.20.a). Using simple primitives (width between two chain lines or the size of the watermark) or more complex features are elaborate and not intuitive to choose. For these reasons, other approaches have been retained and developed.



**Figure 2.20: (a) Watermark Main page. (b) Page that allows the setting of the different values. (c) Result page returned by a query.**

In an attempt to expand the retrieval options, a set of morphological characteristics has been developed and for each characteristic an automatic algorithm has been developed to extract these features from the input images. Total of twelve different features were developed.

Figure 2.21 illustrates the global features used for the retrieval process.



**Figure 2.21: Global features on watermarks. (a) Space between chain lines and the density of laid lines per unit length; (b) Height and width of the watermark; (c) Number of regions; (d) Number and position of junctions; (e) Position of the watermark depending on the chain lines [90].**

The authors have developed two algorithms to facilitate the process of watermarks search and retrieval. The first algorithm is based on computing a circular histogram around the center of gravity of the watermark (Figure 2.22.a). Each bar of the histogram corresponds to the number of pixels present inside a radial segment (Figure 2.22.b).



**Figure 2.22: Graphical representation of the circular histogram. (a) The center of gravity of the watermark is used as the middle point; (b) The number of pixels present inside a radial segment.**

The histograms are computed once for each existing watermark and stored inside the database. Two watermarks are defined as globally similar, if their respective histograms are similar. The similarity measure  $d(H_1, H_2)$  between two histograms is computed as following:

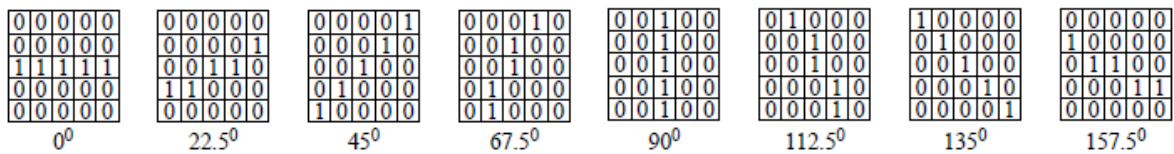
$$d(H_1, H_2) = \sum |H_1[i] - H_2[i]| \quad (2.1)$$

The retrieved watermarks are visually inspected to evaluate the true match.

The second algorithm consists of filtering the input image by eight directional filters.

Figure 2.23 displays the eight filters  $F_j$  used to compute  $G_j$ .

$$G_j(x, y) = I(x, y) \cdot F_j \quad j = 1, 2, \dots, 8 \quad (2.2)$$



**Figure 2.23: Eight filters  $F_j$  used to compute the directional information on the images.**

After this first operation, eight new planes are obtained. We compute the eight  $K(x, y)_j$  planes by taking the highest value from these planes to form an eight dimensional vector  $T_j$  :



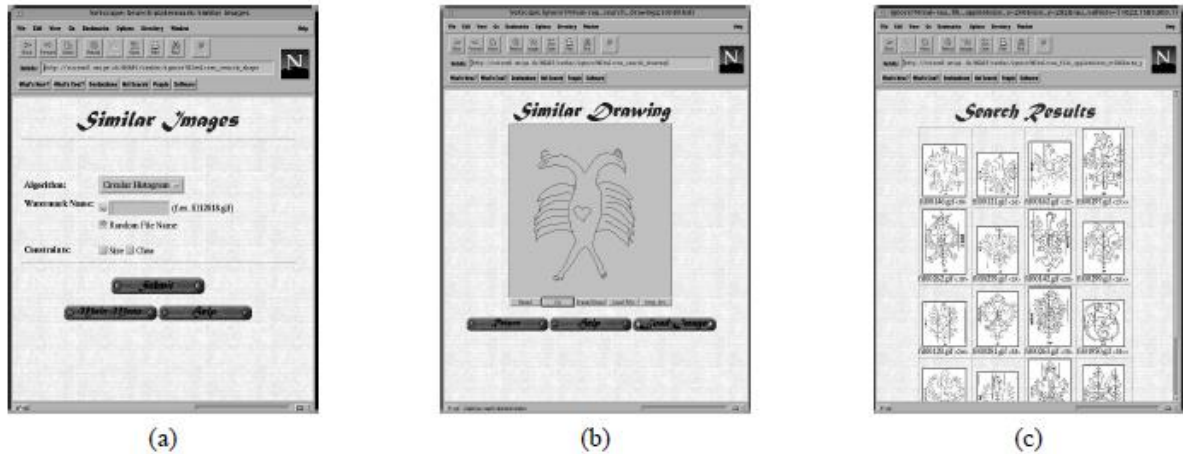
$$K_j(x, y) = \begin{cases} G_j(x, y) & \text{if } G_j(x, y) = \max_j (G_j(x, y)) > 1 \\ 0, & \text{otherwise} \end{cases} \quad (2.3)$$

$$T(j) = \sum_{x,y} K_j(x, y) \quad (2.4)$$

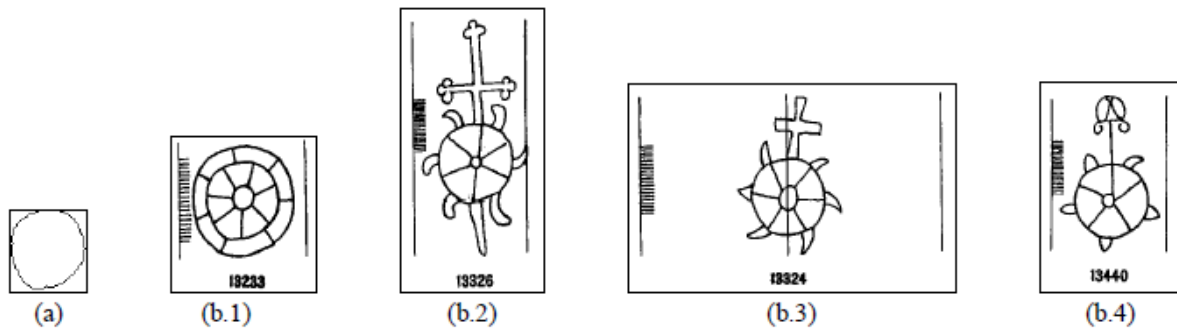
The authors report a typical computing time for shape-based retrieval in the order of 4 seconds per watermark inspected, assuming a 42 x 48 pixels search pattern, a database containing 3,000 watermarks and a 1000 x 1000 x [1..3000] hash table. The authors considered this time as excessive for a blind search in the whole database and they concluded that separating the database in subsets by means of other query criteria is needed to speed up the search process.

As already mentioned, an additional interesting approach in the project is the development of a Java-drawing interface. The interactive interface allows the sketching of a watermark contour directly from the browser (Figure 2.24b). When a historian does not have access to an image of the watermark to retrieve, an approximate shape of the watermark can be drawn. The approximate sketch is then used as the model for retrieval by similarity. With this interface, it becomes possible to load a local image of watermark and to apply some image pre-processing algorithms (for example, erosion or dilation with connectivity four or eight can be applied). A thinning algorithm is present and a module to add small missing parts of the contour of the watermark can be used. Finally, small spots or noise can be automatically removed. After having drawn the model, the histogram is calculated, and the same distance algorithm is used in order to compute the shape representation of the hand-drawn sketch image. The execution time depends on the computer used. The authors report that using Pentium 200 computer (a good tool for that time period) and an image of size 400 by 400 pixels, the algorithm needed

about 30 seconds to extract features and 15 seconds to transfer the data. Figure 2.24 illustrates the drawing interface and Figure 2.25 shows the retrieved of watermarks and shapes.



**Figure 2.24: (a) Main page of the similarity retrieval module; (b) Drawing applet interface to draw a similar shape; (c) Result of similar images.**



**Figure 2.25: Shape-based retrieval. (a) Pattern used as search criterion; (b) Ordered list of the retrieved watermarks: from the most similar ((b.1) to the less similar (b.4) [90].**

The authors have tested their system by taking 120 images from the test-database with a small variation in size, rotation and SNR. The goal was to retrieve the original model without transformation. The reported results include rank of the retrieved images and ROCs on very few tests. If the image was scaled by a factor of 10%, then 40% of the original models are still found in the first position and all models are found in the first 12 images. If 20% of the pixels that compose the watermark are removed, then 85% of the models are found in the first 13 images [90].

Development of watermark descriptors is intended to facilitate their classification in a database and respectively search and retrieval. In his work Zamperoni [94] has proposed a watermark database system in which it is possible to perform watermark image retrieval. He used back-lighting for imaging followed by image processing in order to extract watermarks. First, he removed chain lines using frequency filtering and morphological closing. Then, he used the “top-hat” transform to approximate and subtract the background, followed by contrast enhancement. At the end the images of the two parallel processes are combined using a logical AND operator. The result is finally filtered by a median filter.

In [95], Brunner and Burkhardt use a support vector machine (SVM) with an intersection kernel and leave-one-out tests for classification and retrieval of ancient watermarks. The best reported retrieval results have been achieved with the histogram intersection similarity measure. For the fourteen class problem they obtain a true positive rate of 87%, which they claim “that is better than any earlier attempt”.

In [96], the authors propose a method for trademark image database retrieval based on object shape information that would supplement traditional text-based retrieval systems and report “significant improvement compared to manual retrieval...”. The authors in [97] used three sets of various global moment features and three sets of component-based features. The component-based feature set consists of several shape descriptors which are extracted from various image regions.

Perhaps the most relevant work using image based features for watermarks retrieval is presented by G. Brunner in [98] where further details are given for the work presented in [95]. The author introduced a structure-based feature extraction technique with a method capable of representing the global structure of an image, as well as local perceptual groups and their

connectivity. The advantage of the method is its broad range of applications and its invariance against changes of illumination and similarity transformations. A line segment grouping, based on agglomerative hierarchical clustering, is also used where the segments are extracted with an edge point tracking algorithm. One of the tested applications was the classification and content-based image retrieval of ancient watermark images. The results are presented with averaged class-wise precision/recall graphs that show the discrimination ability of the structure-based feature set. For the classification, the author used a support vector machine with an intersection kernel. The results are obtained by a leave-one-out evaluation and show that for the fourteen class problem, an average true positive rate of 87.41% could be achieved.

The literature review on watermarks research and the selection of references presented in this section clearly indicate the importance of digitization of historic documents and the amount of the work already done in this area. It also illustrates the need for better tools and techniques for imaging and image processing. The examples found in the literature illustrate the obstacles and the limited performance of existing systems for watermarks searching and retrieval.

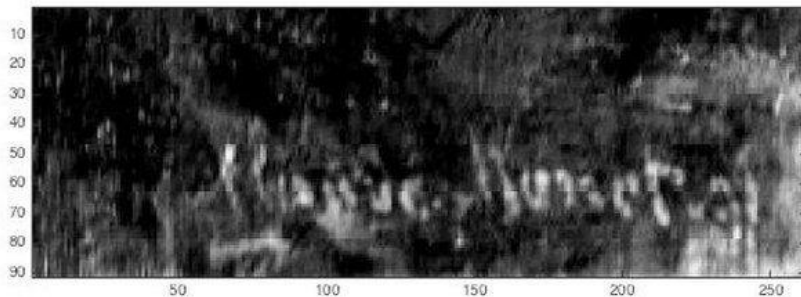
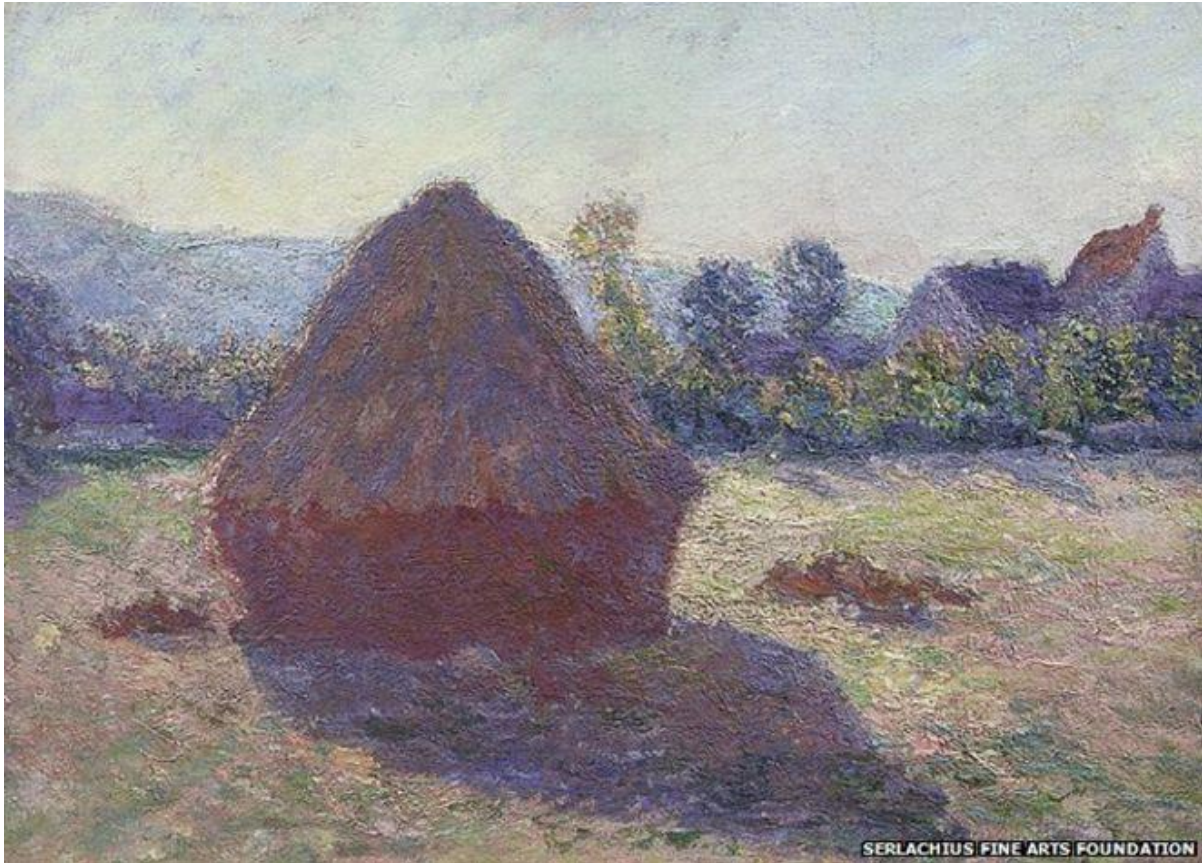
The development of an efficient and accurate search and retrieval method for watermarks is one of the main goals of the current thesis. Before introducing the new method, the next section describes an alternative technique for imaging of historic documents and manuscripts. The method is based on hyper-spectral imaging – a technology originally developed for military monitoring and remote sensing. With the recent technological developments in optical systems, image sensors, cameras, memories and software, hyperspectral imaging becomes more accessible for many other applications.

## **2.4 Hyperspectral imaging of historical documents**

In the last several decades, Hyperspectral Imaging (HSI) rapidly proliferates from its original applications in remote sensing [99] and military purposes [100] to microscopic analysis of biological materials, forensics, chemical detection, and others [101]. Hyperspectral Imaging is a non-destructive optical technique that gathers electromagnetic radiation from across a large portion of the electromagnetic spectrum in many narrow bands to create what is known as the hyperspectral cube. The cube is a collection of monochromatic images (frames) at many specific (center) wavelengths. The XY dimensions of the hyperspectral cube are the image frame coordinates and the third dimension is the spectral wavelengths. Each pixel is presented by a column vector of the electromagnetic energy measured by the sensor at specific wavelengths. This essentially reveals a “spectral signature” of the object of interest by measuring its reflectance, transmission, absorption or luminescence. HSI sensors measure the intensities of these properties at specific wavelengths in regions of the spectrum: ultraviolet, visible, and infrared [102].

The main advantage of HIS is the spectral resolution. One may consider the standard RGB imaging as a broadband signal acquisition that integrates the values over three overlapping bands and inhibits the ability to detect optical properties at specific wavelength. The hyperspectral imaging (HSI) acquires images at multiple (hundreds) narrow wavebands. HSI is a non-contact method, which is especially important concerning the fragile state of some valuable historical documents. HSI provides the possibility to examine spectral segments and spatial layers of the object and analyze elements of their composition. The technique, known as anomaly detection or target recognition, presents the possibility of retrieving text that has been lost through fading, staining, overwriting or other forms of obscuring. HSI offers the prospect of distinguishing different ink types, paint and paper composition, which can be used

to answer questions about the time period and provenance of the documents. Figure 2.26 illustrates the recent signature discovery and authorship conformation of Claude Monet's "A Haystack in the Evening Sun". As the news announcement says, "*RECENART®'s mobile laboratory team visited Serlachius Museums a couple of weeks ago to study some of their paintings. One of the paintings was Monet's painting "Haystack in the Evening Sun" from 1891. The painting did not have a visible signature. ... With hyperspectral imaging the Monet signature was found underneath the paint layers.*" [103].



**Figure 2.26: Claude Monet's A Haystack in the Evening Sun (c 1891) was examined at the University in the Recenart – Research Centre for Art project, Courtesy of Gosta Serlachius [103].**

The application of hyperspectral imaging in the area of historical document examination is a novel and still somewhat specialized area of research, incorporating both machine vision and image processing to provide information about the physical characteristics of questioned documents and historical manuscripts. The recent advancement of related technologies

increased the capabilities of the HSI while decreasing the cost of the systems. Nevertheless, the instrumentation remains specialized and still at a prohibitive cost to become a mainstream. One of the main goals of the CODICES team at UMKC was to develop a HSI system with low cost, portability and ease of operation with respect to imaging of historic documents and books. Before describing the results, a review of some previously developed systems for HSI of documents is presented.

Klein et al. [104] provide an excellent description of the basic concepts, working principles, construction and performance of a HSI device specifically developed for the analysis of historical documents. The device has been used to record the variation of spectral reflectance on a historic seventeenth century map. The instrument also was used to compare the local variation of the yellowness index of reference for bound paper volume and loose sheets. HSI allows the detection and visualization of differences in aging processes of documents, and are particularly useful to monitor the effects of exhibiting and handling on document aging.

Vaarasalo [105] provides a useful discussion on the optical properties of paper, which changes its optical properties with time [106]. As the author indicates, the hyper-spectral images of historical paper, parchment, and other carriers are complex spectral combinations of the optical properties of a collection of materials for writing, printing and drawing as well as the opaque or semi-transparent carrier itself. The materials also have different temporal degradation properties. More recently, De la Rosa and Bautista [107] have discovered that to find the presence and concentration of different colorants or components in the paper, it is only necessary to know the spectra and fluorescence lifetimes at a specific single wavelength (337.1 nm in their case). They indicate that these kinds of measurements could be useful for studying the papers long-term stability and how aging affects it. Additionally, the authors refer to the



particular importance for the preservation of paper-based historical records including the ambient storage conditions as specified by the Committee on Preservation of Historical Records.

Hyperspectral imaging creates a large dataset for analysis purposes, recording vast amounts of information within a wide band of the electromagnetic spectrum with high spectral resolution. This provides a document analyst with information regarding an object's characteristics at specific wavelengths. Other factors such as illumination source, type of emission being examined (e.g. reflectance, transmittance, fluorescence, absorption) also have a bearing on what information is acquired and available to be analyzed. As a result, currently the application of HSI for the examination of historical documents still requires some human interaction or examination to solve document specific questions. In this regard, HSI has to be viewed as a powerful enabling tool. Automation of the systems for image acquisition, including calibration, is important in order to obtain hyperspectral cubes with sufficient and conclusive quality. Automation of the image processing can accelerate the investigation process with the application of modern computational hardware and software.

Even without any form of immediate processing, the hyperspectral cubes are a good method of digitally preserving historical documents and manuscripts. This three dimensional data also provides ample spectral data to perform post processing for more detailed analysis. Technological innovations in hardware (fast CPUs, powerful GPUs, clusters, the cloud) combined with modern processing algorithms (machine learning and parallel processing) substantially facilitate the successful discovery of useful information from the large amount of data. HSI and image processing techniques have been developed and successfully used in the

realm of remote sensing. Their potential in document and art investigation is facilitated by the new technology and the fact that the imaging is performed in a controlled environment.

In the field of historical document examination, Tonazzini et al. [108] have modelled a document as a series of superimposed patterns and have implemented a linear mixture model to describe the relationship between the patterns. Detecting barely perceivable patterns is formulated as the separation of the various patterns in a mixture. A successful discrimination with statistical methods, including independent component analysis (ICA) is used to extract partially hidden texts and for the removal of the effects of the seeping of ink (bleed through) from the reverse paper side [109]. It was found that in general the mixture coefficients are not known, and the separation problem becomes an instance of the blind source separation (BSS) problem [110] where the information in the captured images exists in “mixed signals”.

Principle Component Analysis (PCA) is another dimensionality reduction technique which is often applied to hyperspectral data as a means to isolate, uncover and enhance non-obvious features in document images. PCA is used to condense a high dimensional dataset of interrelated variables to a dataset of lower dimensions retaining as much as possible of the variation between the variables. PCA is used to identify the most representative elements of the data (the principal components), which is of use in hyperspectral image analysis. These principal components are selected and ranked by their relative variance, revealing the most variant representative element to be the most important [111]. In [112], the PCA algorithm has been successfully applied in the analysis of historical parchments in the Baltimore’s Walters Art Museum. The algorithm has aided in the enhancement of faded text and the isolation of different inks. In this work, the first principle component revealed a view similar to that of normal visual inspection. The second component significantly enhanced the legibility of the

partially faded main text, displaying a high contrast between the ink and paper. The third principal component distinguished the individual inks in manuscripts containing more than one type of ink by highlighting differences in the spectral responses at wavelengths outside the visible range.

Hyperspectral image processing can also aid in the reading of documents which have degraded naturally over time by separating the under-written and over-written text to allow enhancing the contrast between them. Analyses at Rochester Institute of Technology and Johns Hopkins University have discovered previously unseen parts of famous Archimedes' palimpsests which are now readable. Another capability particularly important to this thesis, is that hyper-spectral images scanned in the infrared spectrum can be used to better detect and visualize watermarks with techniques developed for remote sensing [113]. The post processing of hyperspectral images has seen much development in the areas of remote sensing of geographical images [114, 115] for purposes as feature extraction, target discrimination, classification and quantification. Borrowing ideas from previous applications and developing them further for the specifics of forensic analysis of historical documents, scholars can achieve very useful results.

Hyperspectral imaging for document analysis is an emerging field of study but has already proved useful in aiding many high profile document examination projects. As a non-invasive, non-destructive imaging technique, HSI is used to capture the optical properties of objects at a series of contiguous wavelength bands. The series of these hyperspectral bands are typically in the hundreds, and contain information from a large portion of the electromagnetic spectrum, encompassing information outside the conventional imaging systems and the natural sensitivity of the human eye. The question of how to digitally represent and preserve the current

state of historical documents is likely the most prominent obstacle encountered by historical analysts and document conservators. Historical documents, with their cultural and sometimes irreplaceable historical values, provide opportunities and challenges to scholars. The solutions most of the time require a multidisciplinary approach and the CODICES team is one very good example.

When light interacts with an object several different phenomena may occur. The incident light may be reflected, absorbed, transmitted or scattered. In general, several of these processes happen simultaneously to a different degree.

**Reflectance** is the process in which a fraction of the incident radiation is reflected by an optical boundary in the direction somewhat opposite (less than  $180^\circ$ , in general) to the incident light. The reflectance spectrum captures the reflectivity of the object as a function of wavelength.

**Absorption** of electromagnetic energy describes the energy changes with transformations to other forms of energy, for example to heat, when light interacts with materials. It relates to the coupling of photons at specific wavelengths with the structure of the radiated object.

**Scattering** occurs when photons are forced to deviate from a straight trajectory by the interactions and non-uniformity in the material through which they pass.

**Transmitted** and **refracted** light is the portion of incident light that passes through a material.

Additional light properties are also important for the imaging process. Light is typically non-coherent, diffused and a mixture of wavelengths. Normally, light (solar radiation, light emitted from incandescent light bulbs, LEDs, etc.) contains many different frequencies, and when it interacts with an object, reflection, absorption and/or transmission at certain

frequencies occurs. All this complexity of interactions could be captured with the collection of a specific wavelengths to form a high resolution spectral characteristic of the object of interest. Another optical phenomenon that can occur and can be measured using HSI equipment is the emission of light as a result of the absorbed energy (fluorescence). There are light sources designed to emit light with specific properties in regard to wavelength, polarization, coherence, etc. The use of such special sources (tunable lasers, for example) provides additional elements of controlled interactions with the object of interest. The photon energy decreases as the wavelength increases (frequency decreases). For that reason the UV light (<400nm) has to be used carefully in order to prevent destructive transformations of the imaged materials. On the other hand, lower wavelengths provide better optical resolution (for example blue light compared to near infrared).

Capturing the spectral response of an object over a broader electromagnetic region provides reach information which can be used to define the composition of the material. Hyperspectral imaging systems capture spectral signatures of materials with high resolution (multiple narrow bands with 10nm at half height, for example). There is no specification for the number of the multiple bands. When they are in the order of tens, the imaging is referred to as *multi-spectral* in comparison with the *hyper-spectral* imaging when the number of narrow bands are in order of hundreds.

#### **2.4.1 Hyperspectral imaging systems for historical documents**

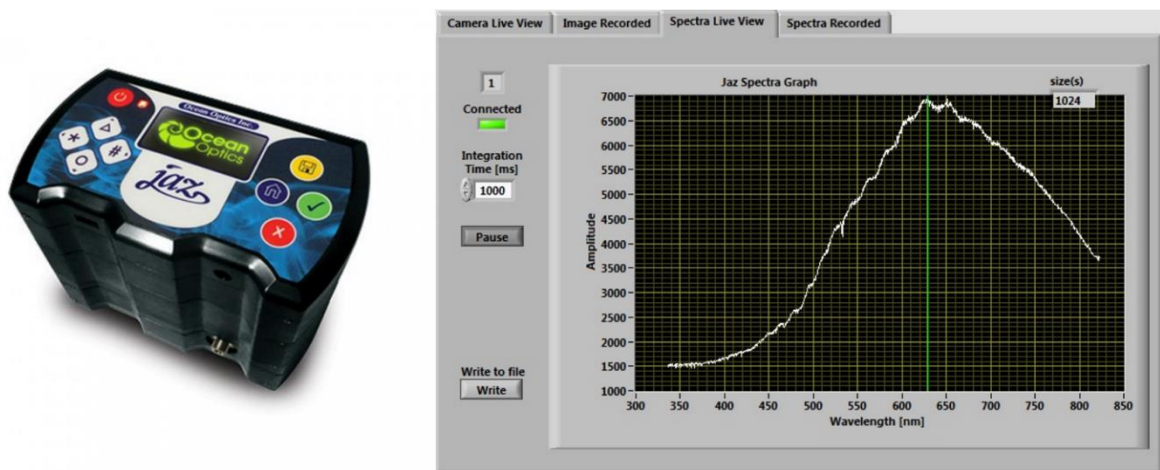
As already indicated, there are many HSI systems commercially available or custom built to specific requirements. The approach of CODICES team at UMKC was to design and build a system at a reasonable cost with readily available components. The goal was to be able to

replicate the system in the future and make it available to multiple organizations. The system had to be compact and easy to use.

Two separate systems were designed. One was intended to scan the page of a book by moving in X-Y plane parallel to the page and at some distance from the surface. The continuous spectrogram is produced for the illuminated area. The repositioning distance is based on the size of the illuminated area. The speed of complete page scan is defined by the scanned surface area and the time specified for scanning at each location. In general, multiple scans are performed for better signal-to-noise ratio.

The second system was designed to take a full-frame size images at predefined wavelengths in the visible and infrared spectra. The two systems are presented next.

For the area small area spectrometry the Jaz EL350 VIS-NIR spectrometer was used (Ocean Optics, Inc., Dunedin, FL, USA). The system was interfaced and controlled with custom LabVIEW (National Instruments, USA) based interface as shown on the right in Figure 2.27.



**Figure 2.27: Ocean Optics Jaz EL350 VIS-NIR spectrometer (left) and LabVIEW based GUI for control and monitoring.**

Jaz is a modular and autonomous module that could be combined to create a community of smart sensing instruments that are unfettered by the limits of traditional optical sensing instrumentation. Jaz is built on a platform that expands to include a light source, rechargeable lithium-ion battery and up to 8 spectrometer modules to make Jaz an analytical instrument easily adaptable for the field, lab or process environment. Jaz can also connect to a computer via the USB port. When connected through a USB 2.0 or 1.1, the spectrometer draws power from the host computer and can be controlled by OceanView software. Based on the drivers provided by Ocean Optics, a custom interface was developed. Figure 2.27 shows the spectrogram of the light source used in the CODICES imaging laboratory. The GUI has tabs to display the Live View or Recorded spectra.

Some of the Jaz features are:

- Sony ILX511B 2048-element linear silicon CCD array detector;
- 14 grating options, UV through shortwave NIR;
- Responsive from 200 nm to 1100 nm;
- Maximum Wavelength range 650 nm;
- Sensitivity of up to 75 photons/count at 400 nm; 41 photons/count at 600 nm;
- An optical resolution of ~0.3 to 10.0 nm Full-Width at Half-Maximum (FWHM);
- Integration times from 870  $\mu$  to 65 seconds;
- EEPROM storage for Wavelength Calibration Coefficients, Linearity Correction

Coefficients, and an optional Absolute Irradiance Calibration;

- SD card data storage;
- Low power consumption of ~ 2.5 W and 8-hour rechargeable lithium-ion battery;
- 16 bit, 3MHz A/D Converter;

- Up to 3 spectrometer channels in a stack for simultaneous multipoint measurement.

The USB connectivity makes Jaz extremely adaptable and easy to operate. The Jaz display can be rotated 180 degrees in either direction with the faceplate specified for either orientation. One can use the specific modules needed for the particular application and to extend the bandwidth of the stacked system. This flexibility and the features which are listed above, makes Jaz a uniquely customizable scientific instrument and eliminating the cost of unnecessary components. That was the reason for selecting Jaz for the design of the HSI system.

The full-frame imaging system was design based on the Cambridge Research & Instrumentation, Inc. (CRi, MA, USA) VariSpec Liquide Crystal Tunable Filters (LCTF, Figure 2.28). The family of filters are solid-state tunable birefringent devices that can be used for both imaging and non-imaging spectral analysis. The modules function like high-quality interference filters, but the wavelengths of light they transmit are electronically tunable and allow for the rapid, vibration-less selection of any wavelength in the visible (VIS) or near-infrared (NIR) region for which the filters have been constructed to operate. Based on its filters, CRi manufactures and sells the Nuance™ Multispectral Imaging System, the R&D 100 Award-winning Maestro™ In-Vivo Fluorescence Imaging System, and the Vectra™ microscope slide analysis system.





**Figure 2.28: Electronically controlled tunable filters.**

VariSpec filters employ electronically controlled liquid crystal elements to select the center wavelength and the transmitted wavelength range while blocking all others. Filter transmittance is sensitive to polarization of the input beam, and is increased by a factor of two if the input beam is polarized along the axis of the input polarizer, as compared with a non-polarized or randomly-polarized beam. VIS models contain an integral, non-removable hot mirror for blocking unwanted NIR light and its heating. The product family includes the following models:

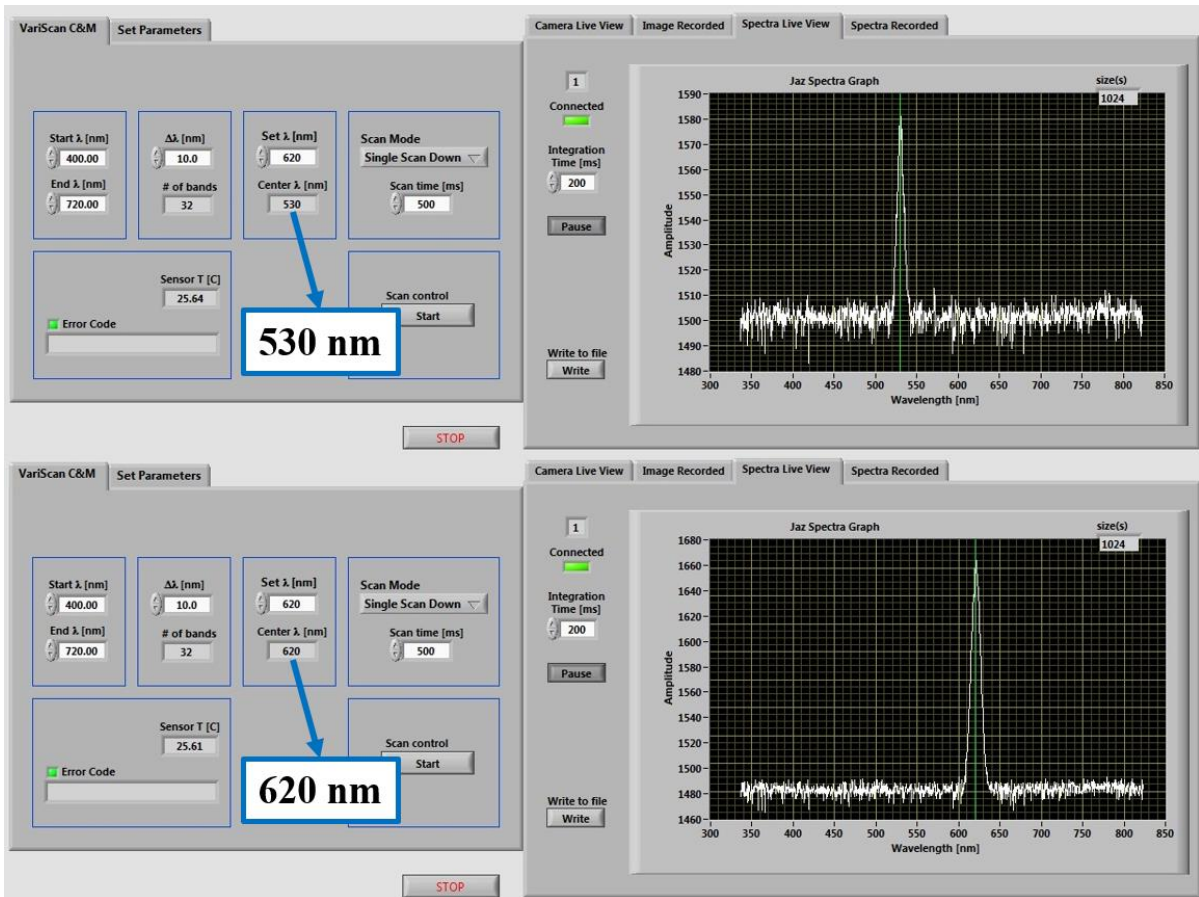
- VIS: Visible-wavelength filters with a wavelength range of 400 nm to 720 nm, bandwidths (FWHM) of 7 nm, 10 nm, or 20 nm, and working apertures of either 22 mm or 35 mm;
- SNIR: Short-wavelength near-infrared filters with a wavelength range of 650 nm to 1100 nm, bandwidths (FWHM) of 7 nm or 10 nm, and a working aperture of 22 mm;

- LNIR: Longer-wavelength near-infrared filters with a wavelength range of 850 nm to 1800 nm, bandwidths (FWHM) of 6 nm or 20 nm, and a working aperture of 22 mm;
- XNIR: Longer-wavelength near-infrared filter with a wavelength range of 1200 nm to 2450 nm, bandwidth (FWHM) of 9 nm, and a working aperture of 22 mm;
- VISR: Visible-wavelength filter with a wavelength range of 480 nm to 750 nm, extraordinarily narrow bandwidth (FWHM) of 0.25 nm (best effort), and a working aperture of 22 mm;
- NIRF: Near-infrared wavelength filter with a wavelength range of 650 nm to 1100 nm, extraordinarily narrow bandwidth (FWHM) of 0.75 nm (best effort), and a working aperture of 22 mm.

In addition to low-power and compact design, other features of VariSpec filters include:

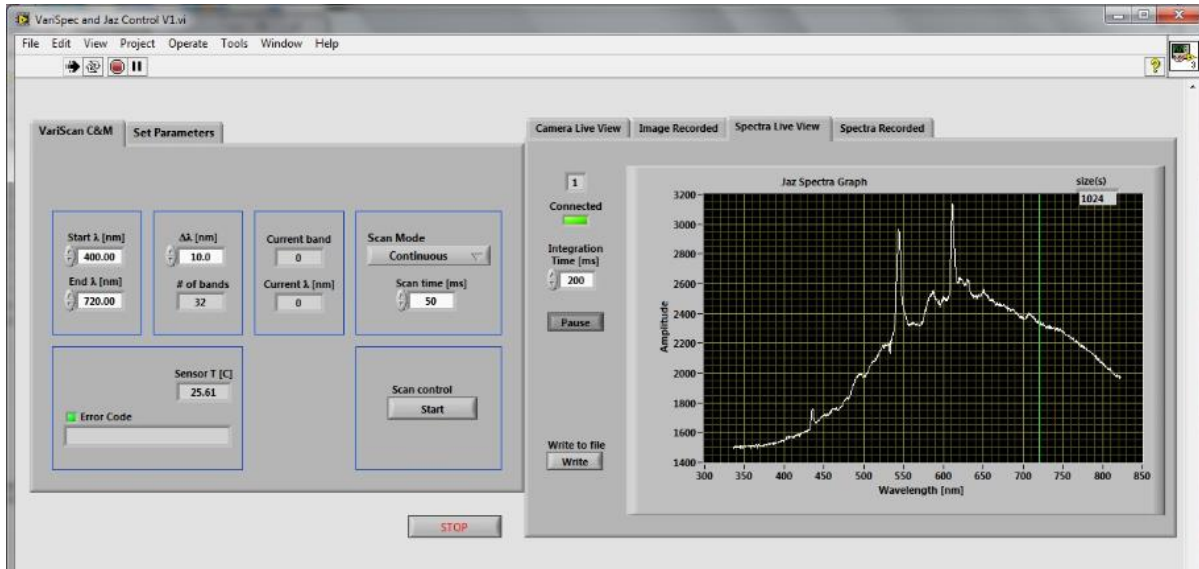
- Ability to tune continuously over hundreds of nanometers in the VIS and NIR;
- Fast, random-access wavelength selection;
- Excellent imaging quality with averaging over multiple scans;
- Solid-state construction, with no moving parts.

Figure 2.29 displays the screen capture of the VariSpec VIS LCT filter integrated with the Jaz VIS module. The top capture is of VariSpec tuned at 530 nm and the bottom at 620 nm. The Jaz's spectral graphs on the right side show the spectral band passing the filter at the exact center frequencies. Several more observations from the graphs: 1) the intensity picks for the two wavelengths are proportional to the illumination light profile given on Figure 2.27; 2) with the integration time set to 200ms, the noise base at 530nm is higher compared to the 620nm indicating lower SNR; 3) at half height the bandwidth of the filter is about 10 nm as specified.



**Figure 2.29. Performance of the integrated Jaz and VariSpec modules.**

For comparison, Figure 2.30 shows the spectrogram of the light in the imaging laboratory when the overhead fluorescent light is on. The picks are common for fluorescent sources and obviously they can severely change the hyperspectral images acquired close to these two wavelengths.



**Figure 2.30: Control panel for VariSpec and Jaz with the spectra of 3200K light source and the overhead room fluorescent light on.**

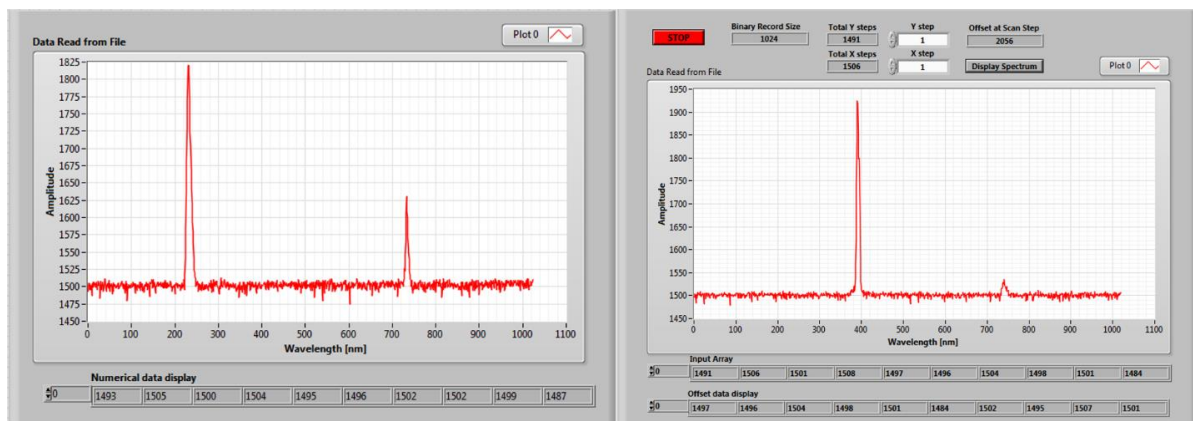
The VariSpec LCTF can be tuned from 400nm to 720nm. In this particular example with 10nm specified bandwidth, there are total of 32 bands. The Jaz acquires spectra from 335nm to 825nm in 1024 data points. The internal temperature of the LCD filters is continuously monitored (25.6° C in this case), which is very important when operating using a light source with NIR spectra. The damage threshold of the filters for reasonably long-term exposure to VIS and NIR energy is 500 mW/cm<sup>2</sup>. Note that the filters absorb the light that is blocked. VIS filters have a near-infrared hot mirror to reflect some unwanted near-infrared light, but longer wavelength near-infrared light, such as thermal energy, can damage the filters.

There are other technologies for electronic tunable filters. Acousto-Optic Tunable Filters (AOTF) use a crystal in which vibrational waves at radio frequencies allow a single wavelength of light to pass. The wavelength of light is a function of the frequency of the radio waves applied to the crystal. Transmission can be high, but without the superior out-of-band transmittance performance of an LCTF. There is also a so-called blur effect that can

compromise the spatial imaging quality of the filter. Field-of-view is narrower than with VariSpec LCTFs because of the requirements of the AOTF design. Tuning speeds can be as fast as 20 microseconds because all that is required is that the radio waves penetrate the crystal.

There are several technical characteristics of the VariSpec LCTF which are not the most optimal. Only the VIS filters come in 22 mm and 35 mm apertures. All others are available in 22 mm apertures only. The filters need from 50ms to 150ms to tune from one center wavelength to another and they linearly polarize the passing light. These constraints are not detrimental for the design of the HSI system but nevertheless limit the design flexibility.

To complete the tools available for small area hyperspectral imaging, several additional executable programs were developed. They provide the functionality to read recorded spectral lines in binary format and to read concatenated spectra when multiple scans are recorded consecutively in the XY plane. Examples of the interface are shown in Figure 2.31.



**Figure 2.31: Spectra reader for binary and concatenated binary spectral records.**

For the full-frame imaging system, the camera could be any commercially available camera that meet the desired image resolution. Considerations are remote control, compatibility of the optics with the selected LCT filters, raw image capture and export, possibility to remove the infrared (“hot”) filter on the sensor of the camera, and others. The Canon, Nikon and several other popular camera families were considered. Canon 5Ti was selected. In this case the

selection was driven by the ability to integrate the camera for a complete remote control with the same GUI that is used to control the LCT filters. The Canon T5i is shown in Figure 2.32 with the VariSpec filter attached.



**Figure 2.32: The Canon EOS T5i camera with the VariSpec VIS LCT filter attached.**

The manufacturers of most modern instruments provide a software development kit (SDK) for one or more programming languages. To expedite the development process, the drivers for LabVIEW were used to develop the control and monitoring GUI which is displayed in Figure 2.33.

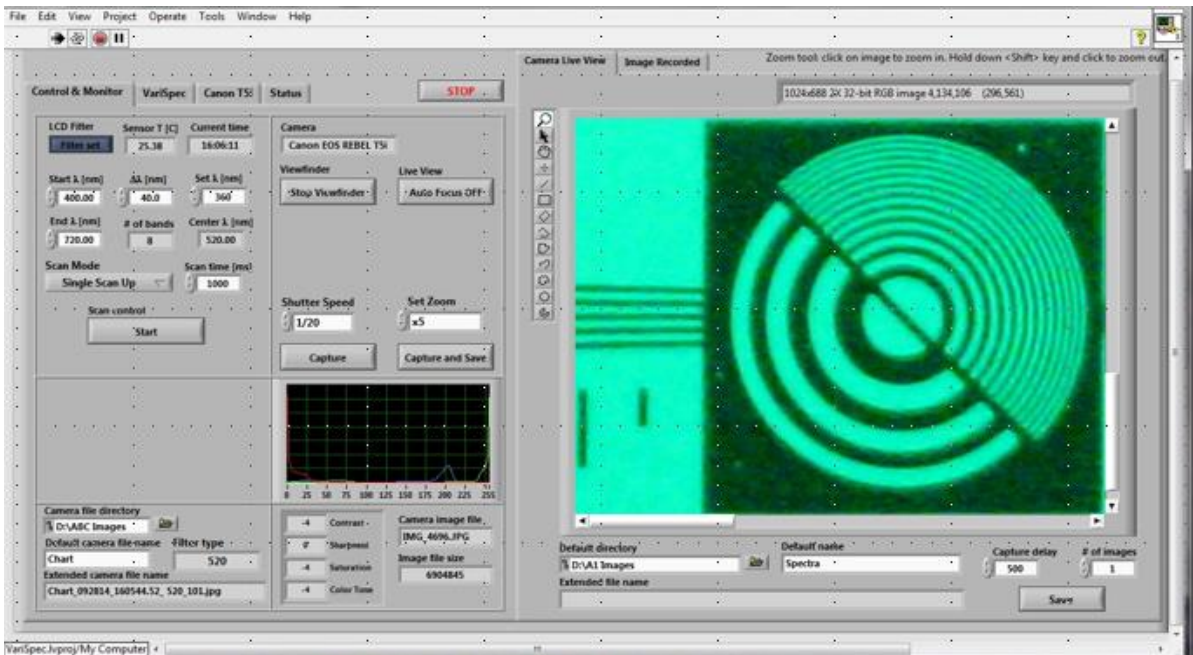


Figure 2.33a: Optical performance test at 520 nm wavelength.

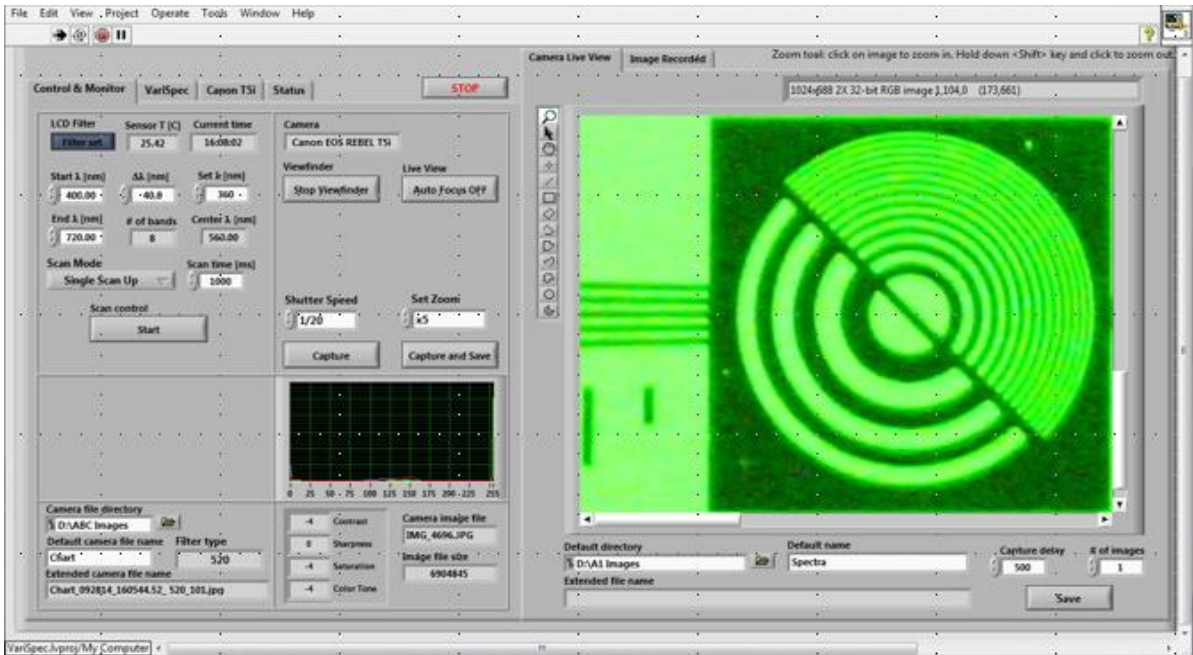
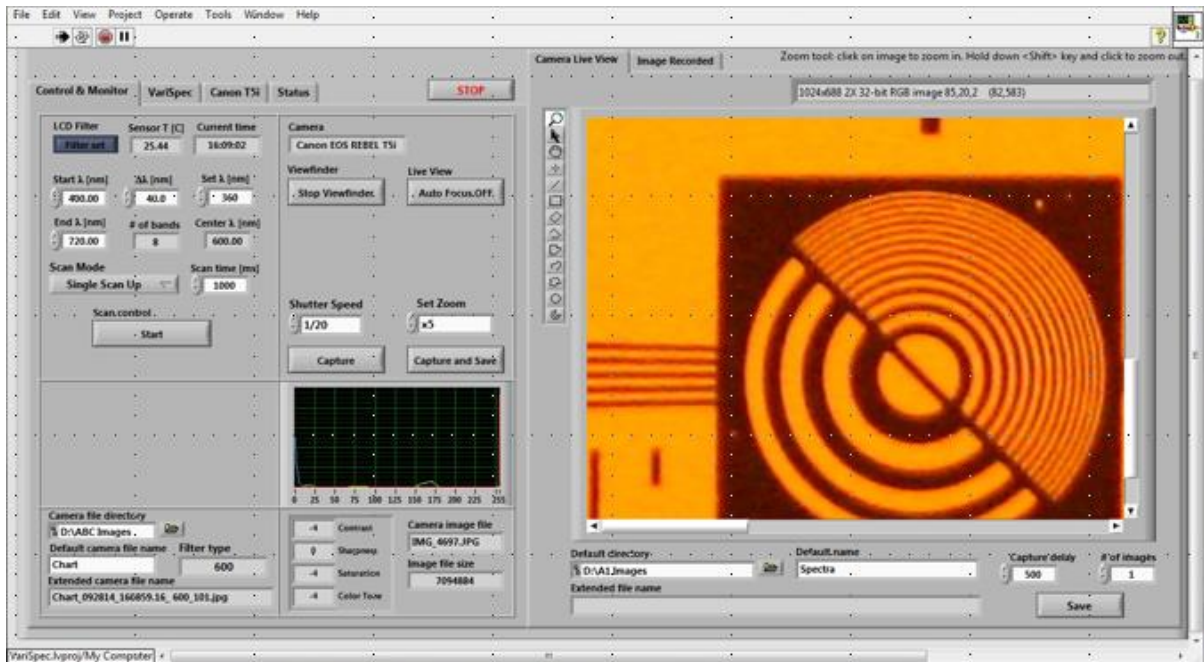


Figure 2.33b: Optical performance test at 560 nm wavelength.



**Figure 2.33c: Optical performance test at 600 nm wavelength.**

The main tab is for acquisition control and monitoring. Three additional tabs provide access for remote setting of the filter, the camera, and for the overall status of the system. Figure 2.33a, 2.33b, and 2.33c display screenshots of the image capturing of a standard test chart at three different wavelengths. With the use of the zooming tool (located next to the top left corner of the image viewer), a magnified portion of the captured chart is displayed to demonstrate the spatial resolution for the three different wavelengths. The GUI has several tabs for image viewing, capturing and storing, as well as control setting for the camera and the filter. The user can observe in real time selected image quality measures (contrast, intensity, saturation, color components). The user may operate the system in sequential target viewing, adjustments, and single image capture. In automated mode, a set of images is captured and stored. The number depend on the total bandwidth (start and stop wavelength) and the step size between the consecutive center frequencies.



Figure 2.34 is snapshot of the capturing process of *Summa Theologica* in the CODICES imaging laboratory.



**Figure 2.34: CODICES imaging of *Summa Theologica*.**

The initial test of the integrated system was performed with a set of images captured with the Canon 5Ti camera and the VariSpec VIS LCT filter at center frequencies every 10 nm, starting from 440 nm and finishing at 700 nm, for a total of 28 images. The region of interest (ROI), in this particular test example, was the lower left corner of the vertical page of *Summa*. The ROI contains a hand drawing in color. Figure 2.35 displays the RGB image of the first page of *Summa Theologica* volume one without any filtering. Figure 2.36 displays the RGB of a small area from the same page and three frames from the blue, green and red spectra captured with the VariSpec LCTF.

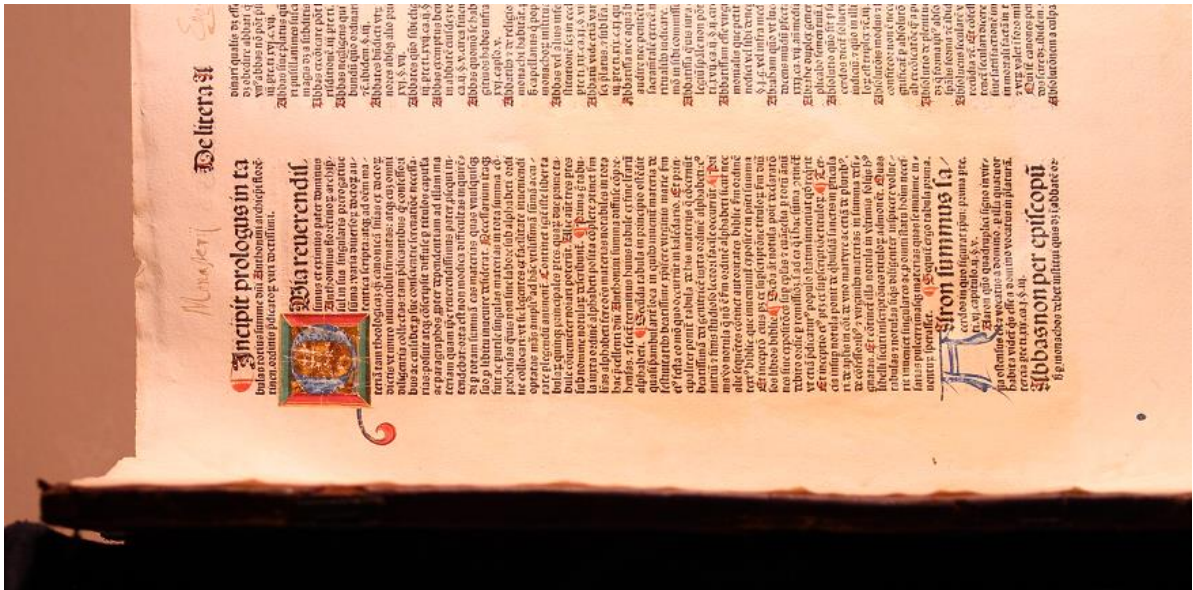


Figure 2.35: The RGB jpeg image of *Summa Theologica*, Volume 1, page 1.

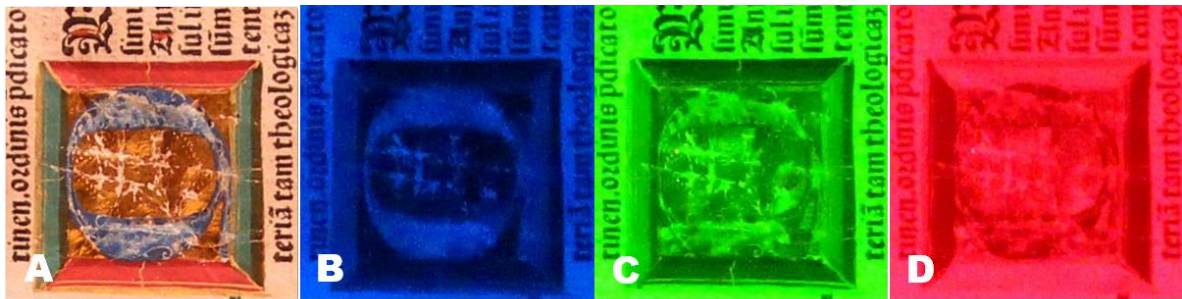
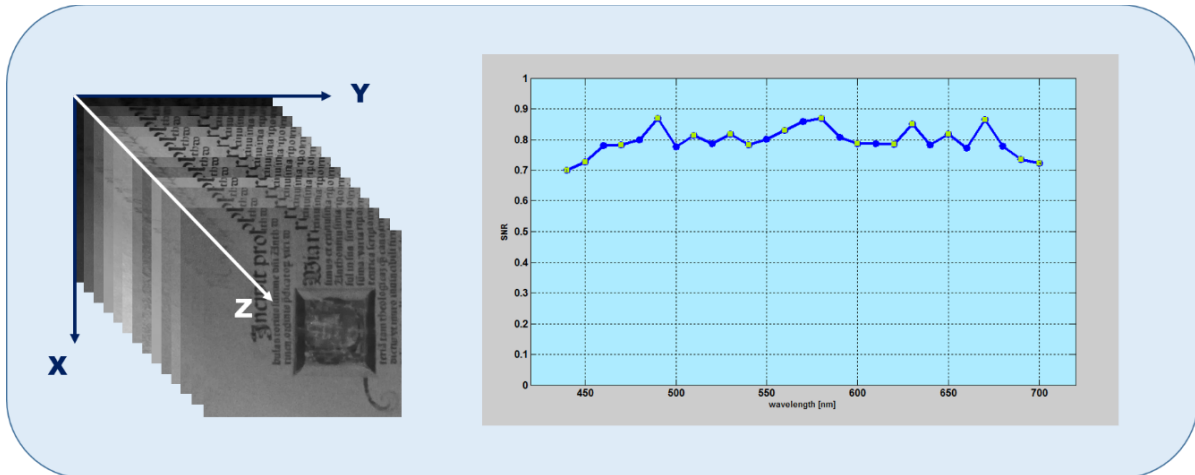


Figure 2.36: The jpeg images from page 1 of *Summa Theologica Prima*. Standard RGB image without filtering (A). The same ROI with VariSpec tuned at 450nm (B), 550nm (C), and 650nm (D).

The subset of hyperspectral images is used to form a mini hyperspectral cube (Figure 2.37, left). The RGB images were transformed to gray-scale and the relative SNR of each image was calculated to evaluate the imaging quality across the whole bandwidth (Figure 2.37, right). Even this simple preliminary evaluation indicates the roll-off of the image quality at the end of the spectrum. The spectral intensity profile of the light source combines with transmission efficiency of the LCTF to further emphasize this effect. This prompts the necessity for real time image quality evaluation during the acquisition process. Or, postprocessing equalization with respect to the illumination source and the spectral characteristic of the filter. When the

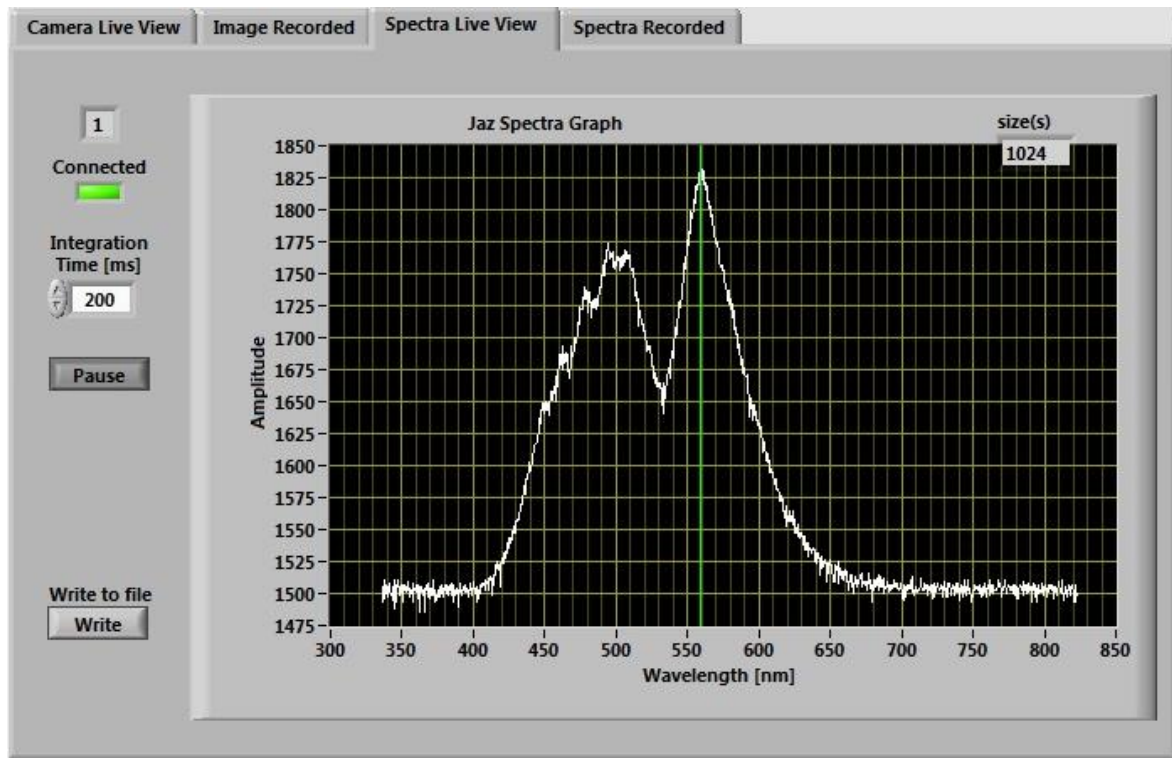
combined transfer function of illumination source/filter is known, it can be used to adjust the camera settings during acquisition. Adjustments for intensity is one aspect. Focus adjustment is also necessary to compensate for the wavelength dependence. Image acquisition in the controlled laboratory setting is an advantage and also the possibility to use an optical system with sufficient depth of field.



**Figure 2.37: Grayscale images from 400nm to 700nm and the calculated relative SNR.**

In codicology, the two sides of each physical sheet or *folium* are referred to as *folium rectum* and *folium versum*. Translating the Latin *rectō foliō* and *versō folio*, it means "on the right side of the page" and "on the turned side of the page", respectively. The short terms *recto* and *verso* are used referring to the text written on the "front" and "back" sides of a page in a bound set such as a codex, book, broadsheet, or pamphlet. As mentioned previously, digitization of some historical documents and manuscripts benefits from backlight imaging which facilitates imaging of the watermarks and other paper artefacts. It is also helpful for digital processing involving text layers removal and subtraction of reflected and transmitted light images. For a backlight imaging, the CODICES team used a tin flat light source and its spectral profile was acquired with the Jaz spectrometer and is displayed in Figure 2.38. The composed "white" (relatively broad) color has two peaks in the upper blue (500nm) and green

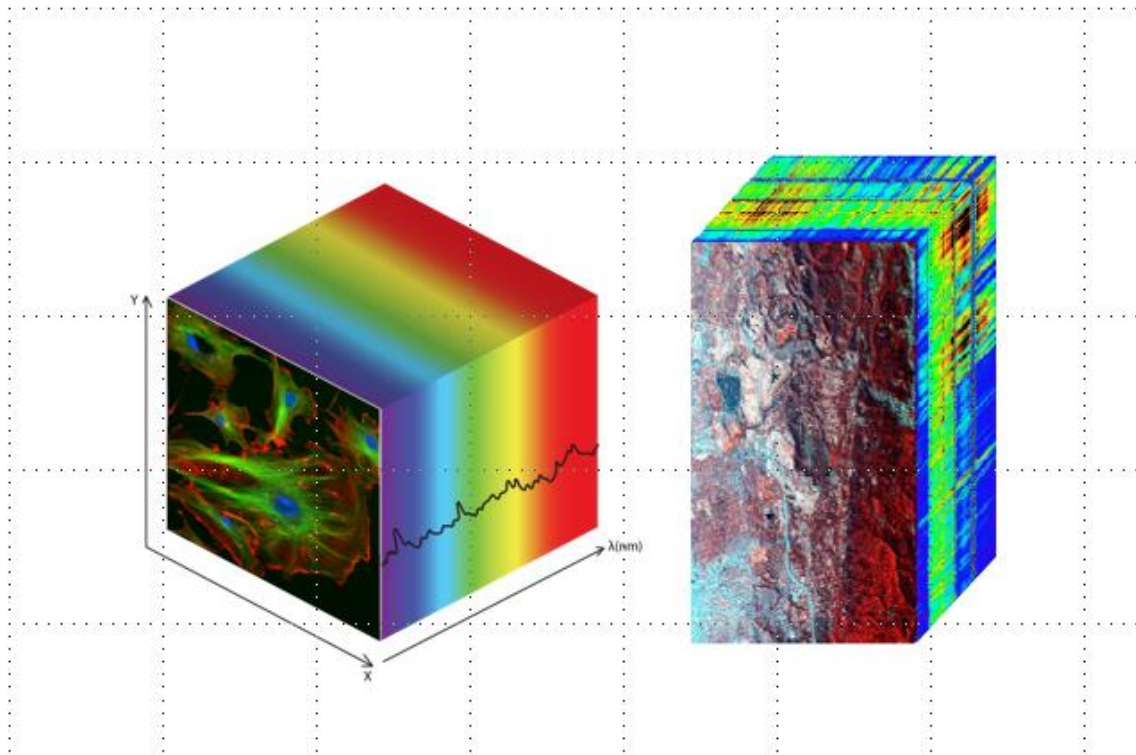
(560nm) regions. Based on the spectral profile, this source cannot be used for hyperspectral imaging or spectral profile calibration needs to be used.



**Figure 2.38: Spectrogram of “white” flat light source.**

#### **2.4.2 Processing of the hyperspectral cube**

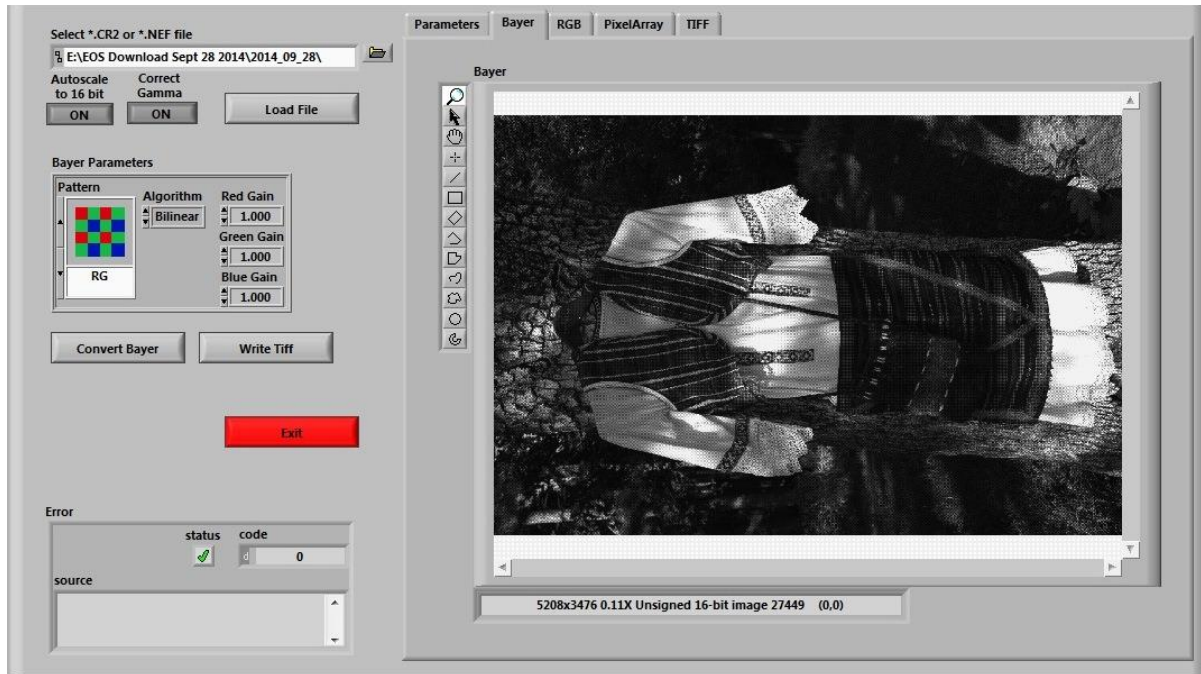
The term hyper-spectral cube (HSC) is not necessarily the most accurate name for the gigabytes of hyperspectral images. Figure 2.39 displays two HSCs. One on the left is from HSI of biological cells and resembles the stereo-metric definition of cube. The one on the right is the HSC acquired with the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) and has 224 bands from Jasper Ridge in California. The very fact is that the x and y dimension of the “cube” are number of pixels (sensor dependent), while the third dimension is a wavelength or the number of frequencies at which an image is acquired.



**Figure 2.39: Two examples of typical hyperspectral cubes.**

During the work on the *Summa Theologica* project, several software tools were developed to facilitate the exploration and processing of the hyperspectral cube (HSC). The images were stored in CR2 (raw) and JPEG (compressed) format. CR2 (Canon Raw Version 2) files refer to the raw image format used by Canon cameras. It stores information directly from the sensor to the storage device. The image cannot be easily viewed like a JPG or PNG format images. When stored in RAW formats, the images can be edited in much more detail in post processing without any format-based loss of quality. CR2 images are not yet ready for printing and cannot be edited with a bitmap editing program. For the CODICES project, software was design to facilitate the conversion from CR2 to gray scale images based on the RGB Bayer pattern of the imaging camera. Figure 2.40 illustrates the interactive exploration for the conversion based on the Bayer's pattern, band gain, gamma correction and the type of output image. This particular image was taken outdoors in direct Sun light. It is an image of folklore female

apparel from the Northern region of Bulgaria which is rich in color and patterns. The overall goal is to explore and demonstrate the transformation capabilities with maximum possible fidelity. While the recorded jpeg images can be used for data exploration and discovery, the high resolution camera data may be used for a detail search and an in depth description of artefacts.

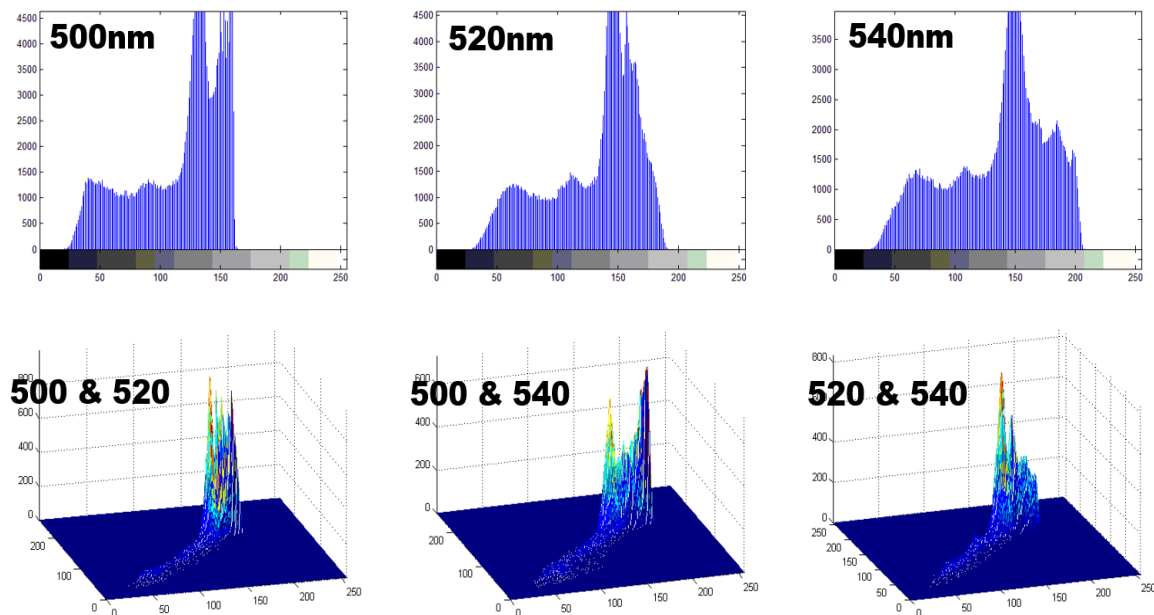


**Figure 2.40: GUI for interactive conversion of CR2 files to TIFF, PNG, BMP, and other image formats.**

The development of preprocessing and evaluation tools with LabVIEW has provided an accelerated development based on the Visual Processing toolbox. Many additional tools are available as commercial products and open source. The processing algorithms for the work reported in this thesis were developed in Matlab (MatWorks, USA). Some exploratory work was conducted in Python, R, and LabVIEW.

The HSI produces a vast amount of data. Traditionally named the hyperspectral cube (HSC) it has GBs of data. The goal is to transform this data to meaningful and valuable information. Given the limited perceptual capability of the human visual system with respect

to dimensions of the data, most of the time the first task that is performed on the HSC is dimensionality reduction. This is especially important during the development cycle of an algorithm and typically aims at visualization in a semi-automatic and iterative method. Most HSI systems have proprietary software for image exploration, dimensionality reduction, and visualization. Other options include a spectrogram for a selected pixel or an average spectrogram for a ROI. A different method is to compute the joint histogram of the images in the HSC and rank the pairs with highest variability. For example, Figure 2.41 displays the individual and some of the joint histogram for three of the *Summa* images.



**Figure 2.41: Individual and joint image histograms for three wavelengths.**

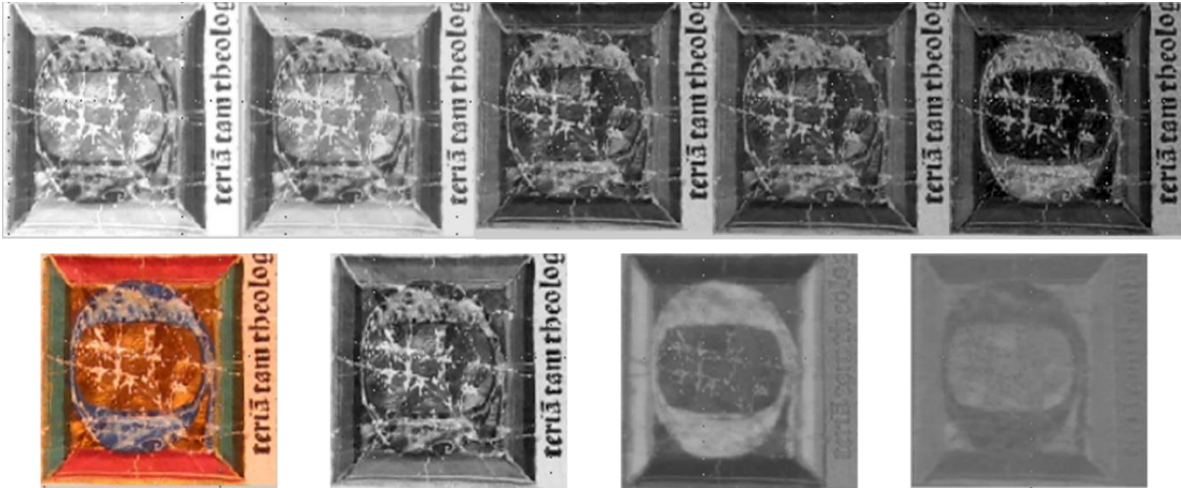
The pixel, ROI, and image histogram approaches provide the user with ability to discover point of interest (abnormality in the spectra) which in general are indicative of “something different from the surrounding area”. In the case of remote sensing and reconnaissance the terminology used is abnormality detection or target recognition. The terminology is suitable for the application of HSI as a forensic tool in the study of historic documents. The terminology

depends on the specific application: spectral un-mixing, blind source separation, feature extraction, and others.

There are several classical methods for dimensionality reduction which were already described briefly. Principal component analysis (PCA), Independent component analysis (ICA) and Knn clustering are frequently used and several additional methods were developed for remote sensing applications.

Principal Components Analysis (PCA) is the most widely used linear-dimension method based on second order statistics. PCA is also known as the Karhunen-Loeve transform, singular value decomposition (SVD), empirical orthogonal function (EOF), and Hotelling transform. PCA is a mathematical transformation that results in the simplification of large data sets by processing a number of correlated variables into a smaller number of uncorrelated variables called principal components [116]. The basic applications of PCA, as applied to hyperspectral images, are dimensionality reduction of the data (data compression) and information extraction, associated with segregation of noise components. This is done by finding a new set of orthogonal axes that have their origin at the data mean and are rotated to a new coordinate system so that the spectral variability is maximized. The resulting principal component bands are linear combinations of the original spectral bands and are uncorrelated. PCA is a family of techniques: the basic PCA is enhanced and developed to kernel PCA, Maximum Noise Fraction (MNF), Noise Adjusted Principal Components (NAPC), and others. Figure 2.42 illustrates the basic PCA applied to hyperspectral images of *Summa Theologica*. The first three principal components are shown next to the RGB image of the same ROI on the bottom row in Figure 2.42:





**Figure 2.42: PCA applied to hyperspectral images of *Summa Theologica* (top row). The first three principal components are shown next to the RGB image of the ROI.**

The minimum (or maximum) noise fraction (MNF) is a second algorithm of the PCA family techniques developed and used in the remote hyperspectral imaging. It was introduced by Green, Berman, Switzer, and Graig in 1988 as a method that takes into account sensor noise [117]. It consists of two consecutive PCA transformations and requires an estimate of the signal and noise covariance matrices. While the basic PCA considers only the variances of each principal component and assumes that noise is isotropic (“white”), the real sensors’ noise is typically not isotropic. Early imaging sensors operated with an elevated noise levels and the images with monotonically declining variance, produced by the PCA, were not always images of decreasing quality. The MNF transform orders the images in terms of the SNR metric, thus ordering them based on image quality. Modern imaging sensors have better characteristics and in controlled environments could be used without active cooling to reduce the thermal noise. This removes the need to apply the MNF.

Because the MNF transform orders the component images based on their quality, it is invariant to scale changes. The component images are ordered based on SNR instead of variance like in basic PCA. MNF transformation is also called the Noise Adjusted Principal

Components transform. It is equivalent to sequentially transforming the data to a coordinate system in which the noise covariance matrix is the identity matrix followed by a principal components transformation.

Dimension reduction is not an easy task and many methods have been proposed to produce the intrinsic dimensionality of the data set. The goal is to retain the minimum number of free variables needed to model the data without or with minimal loss [118]. The reviewed literature describes four methods for reducing the dimensionality of multispectral and hyperspectral image data.

Cumulative variability is the straightforward method for dimensionality reduction. The criterion is that the first components that account for at least 90% of the total variability are retained. It is estimated that these components capture the useful information of the scene.

In 1966, Cattell [119] has introduced the *scree* test. It is a graphical method in which the eigenvalues are plotted versus the ordered principal components. The sharp drop from a high initial values indicates the components that should be retained. The low values (low variance elements) in the tail of the curve represent only the random variability in the data.

The Kaiser criterion [120] for determining the intrinsic dimension of the data is based on the correlation matrix. The transform bands are evaluated and those with a variance greater than or equal to one are retained. Transform bands with a variance greater than or equal to one contain at least as much information as the original.

Continuous Significant Dimensionality (CSD) for dimension reduction was proposed by Yury [121]. It is defined by the formula

$$CSD = \sum_{j=1}^k \min(\lambda_j, 1) \quad (2.5)$$

where  $\lambda_1, \lambda_2, \dots, \lambda_k$  are the eigenvalues of the correlation matrix. The CSD value indicates the number of principal components that should be retained.

In general, the dimensionality reduction of hyperspectral data is achieved with different success and can produce quite different results due to the ill-conditioning of the dimension estimation problem [118].

Independent Component Analysis (ICA) has received attention because of its wide range of potential applications [122 – 128]. The goal of ICA is to recover independent sources given only sensor observations that are unknown linear mixtures of the unobserved independent source signals. In contrast to correlation-based transformations such as Principal Component Analysis (PCA), ICA algorithm includes higher (beyond second order) statistics. ICA not only de-correlates the signals (2nd-order statistics) but also reduces higher-order statistical dependencies, attempting to make the signals as independent as possible. ICA is designed not to search the principal components, which allows to represent the maximum of the return dispersion, but the more independent factors which can linearly generate the returns. Thus, ICA seems to be an attractive candidate for dimensionality reduction.

ICA is based on a linear decomposition of observed data into statistically independent components. Given an observation:

$$\mathbf{X} = \mathbf{A}\mathbf{S} \quad (2.6)$$

where  $\mathbf{X}$  is the vector representation of the recorded signals and  $\mathbf{A}$  is a scalar matrix of the coefficients of the source signals represented by the vector  $\mathbf{S}$ , ICA finds a separating matrix  $\mathbf{W}$  such that:

$$\mathbf{Y} = \mathbf{W}\mathbf{X} = \mathbf{W}\mathbf{A}\mathbf{S} \quad (2.7)$$

where  $\mathbf{Y}$  is a vector of independent components. ICA looks for a linear representation that minimizes an objective function (maximizes a given non-Gaussian measure). A commonly used function in ICA algorithms is the mutual information of vector  $\mathbf{Y}$ :

$$I(\mathbf{Y}, \mathbf{W}) = H(\mathbf{Y}\mathbf{i}) - H(\mathbf{Y}) \quad (2.8)$$

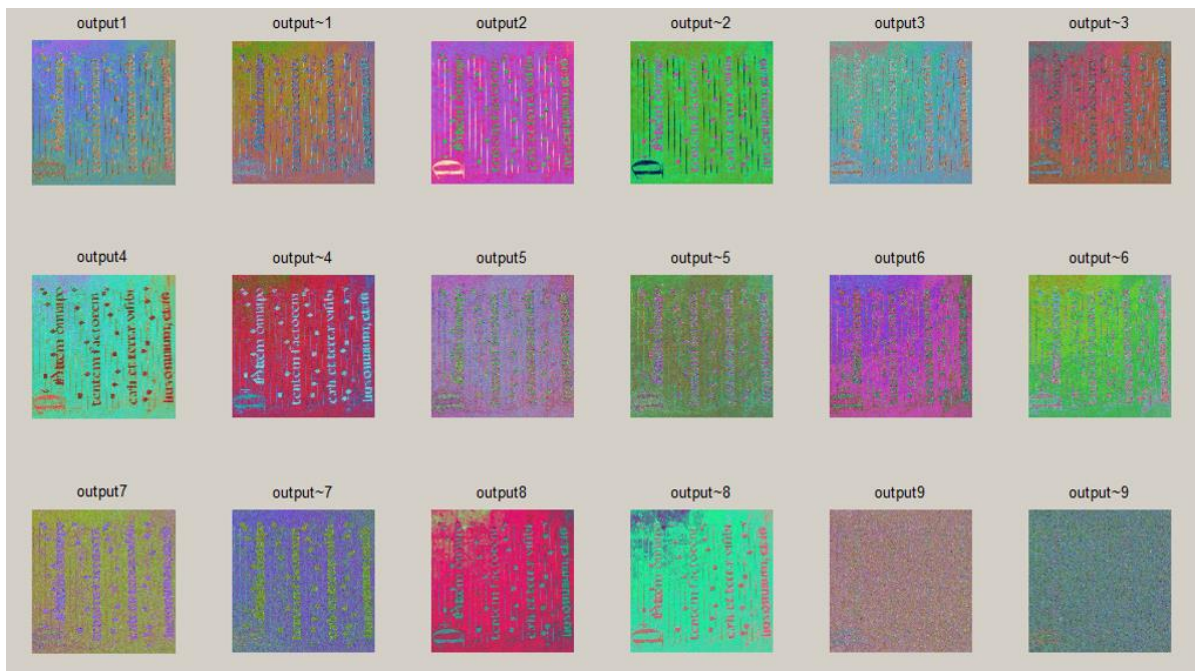
where  $H(\mathbf{Y}\mathbf{i})$  and  $H(\mathbf{Y})$  are the entropy of the random variable  $\mathbf{Y}\mathbf{i}$  and of random vector  $\mathbf{Y}$ , respectively. More details about the general framework of ICA can be found in [129]. Different implementations of ICA are listed in Table 2.1.

**Table 2.1: Different implementations of ICA algorithms.**

1. SANG	(Self Adaptive Natural Gradient with nonholonomic constraints)
2. NG-FICA	(Natural Gradient - Flexible ICA)
3. NG-OL	(On-line adaptive Natural Gradient)
4. EVD 24	(Algorithm based on symmetric EVD 24)
5. ERICA	(Extended Robust ICA - based on Cumulants)
6. SIMBEC	(SIMultaneous Blind Extraction using Cumulants)
7. UNICA	(Unbiased quasi Newton algorithm for ICA)
8. JADEop	(HOS Joint Approximate Diagonalization of Eigen matrices optimized)
9. FPICA	(Fixed-Point ICA)
10. PEARSON opt.	(Pearson system optimized)
11. BSE-C	(Blind Source Extraction - optimization of Cumulants)
12. BSE-K	(Blind Source Extraction - optimization of Kurtosis)
13. SOBI-RO	(SOBI with Robust Orthogonalization)
14. BSS SVD	(BSS SOS algorithm based on SVD)
15. SONS	(Second Order Nonstationary Source Separation)

Independent component analysis (ICA) is a very popular (and classical) method that has shown success in applications for feature extraction and unsupervised recognition. There are many enhancements and implementations of ICA. Some of the algorithms are listed in Table

2.1. In general, the statistical independence of sources in hyperspectral images cannot be verified. However, it is a fact that many of the historic documents have been rewritten and corrected multiple times. This is especially true for manuscripts on parchments and other expensive writing surfaces. Also the watermarks and the writings on the paper are by default separate, statistically independent sources. This logic makes ICA a suitable tool in the study of historic documents for sources un-mixing and artefact discovery. The typical use of ICA for dimensionality reduction of the HSC is studied as hyperspectral data classification. Experimental results shown on Figure 2.43 indicate the capability and effectiveness of ICA when used for hyperspectral image pre-processing. A Matlab implementation of the first algorithm of Table 2.1 (highlighted in blue) is used for this example.



**Figure 2.43: ICA algorithm outputs. Images from Adair Chant Book**

There are many other methods for processing the HSC. The number of new algorithms continue to grow and the performance results continue to be improved. The ability to successfully process the images of the HSC and obtain the useful information is extremely

important and the basis for successful outcomes. However, as already mentioned, the acquisition and processing of the HSC is not the main focus of this thesis. The reference materials and demonstrated examples are given for completeness of the systematic approach to the study of historical documents and manuscripts. Some of the referenced methods have multiple application available in the public domain or as open source for academic research.

One important consideration during hyperspectral imaging is to maintain a spatial alignment during the acquisition process. Remote control of camera and filter helps to avoid any mechanical disturbances during a complete acquisition of the HSC. The ability to align images is very useful. Table 2.2 shows some rigid, non-rigid, and optical flow algorithms for image alignment.

**Table 2.2: Rigid, non-rigid and optical flow alignment algorithms.**

1. Lucas-Kanade alignment (forward-additive, multi-resolution)
2. Lucas-Kanade alignment (inverse-compositional, multi-resolution)
3. ECC image alignment (forward-additive, multi-resolution)
4. ECC image alignment (inverse-compositional, multi-resolution)
5. Dual Inverse Compositional alignment algorithm
6. Affine transform from point correspondences
7. Euclidean transform from point correspondences
8. Similarity transform from point correspondences
9. Homography from point correspondences
10. Translation from point correspondences
11. Inverse (backward) warp of image
12. RANSAC-based transform estimation
13. Modify the warp for half- or double-size resolution
14. Nonrigid registration (flow-based) by SIFTflow algorithm
15. Efficient optical flow by SimpleFlow algorithm
16. Nonrigid demon registration algorithm

Random sample consensus, or RANSAC, is an iterative method for estimating a mathematical model from a data set that contains outliers. The RANSAC algorithm works by identifying the outliers in a data set and estimating the desired model using data that does not contain outliers. Figure 2.44 shows an example of RANSAC filtered correspondences for

hyperspectral images of an old music chant. Figure 2.45 shows the error comparison with the Least-square error method.

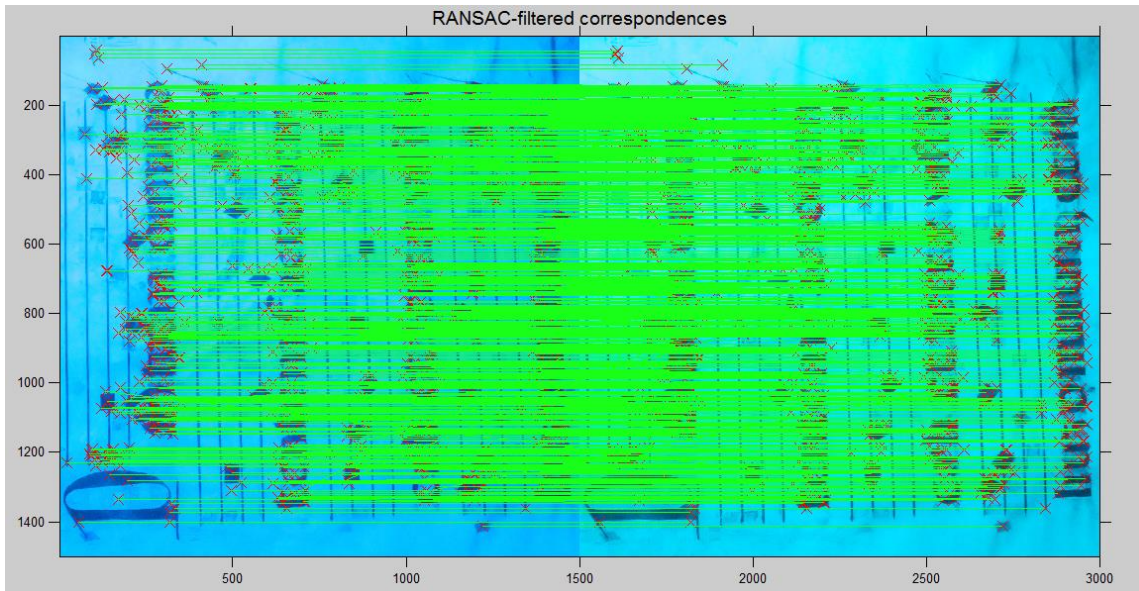


Figure 2.44: RANSAC alignment of images – point correspondences.

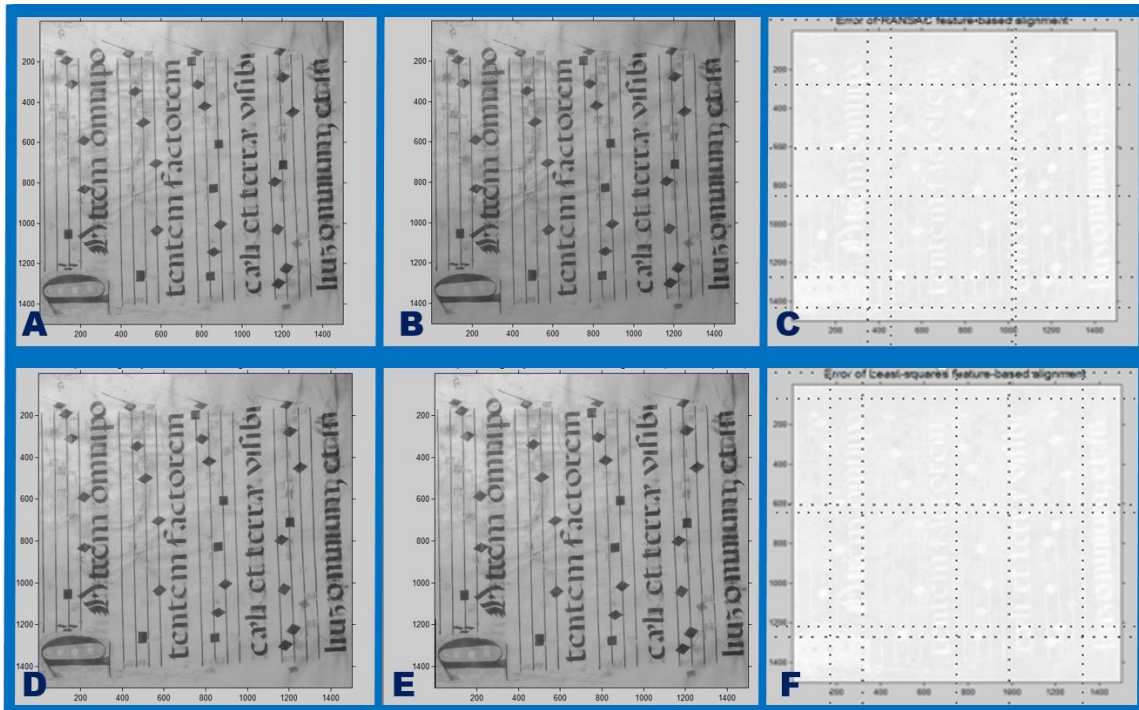
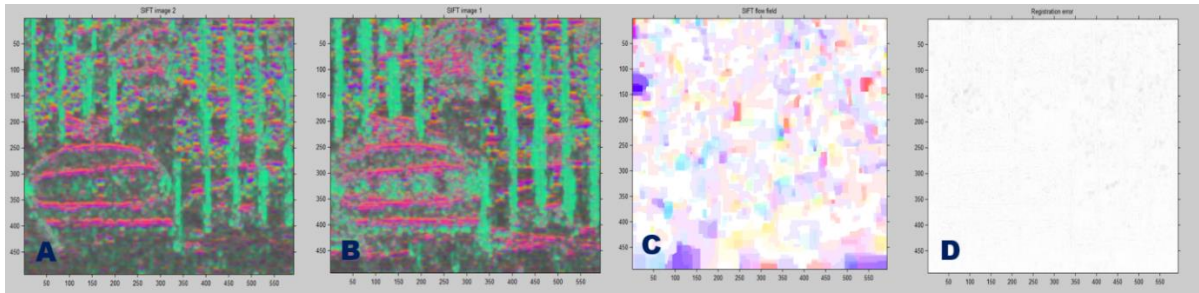


Figure 2.45: Two images (A and B), the error of RANSAC alignment (C), the warped image by features based alignments with RANSAC (D), warped image by features based alignments with Least-squares (E), , and error of Least-squares alignment (F).

Scale-invariant feature transform (SIFT) is an algorithm used in computer vision to detect and describe local features in images. The algorithm was published by David Lowe in 1999. Figure 2.46 shows an example of image alignment using the SIFT transform. The two images are A and B, the SIFT flow field (C) and the registration error (D).



**Figure 2.46: Images alignment with SIFT transform: images (A and B), SIFT field flow (C), and the registration error (D).**

For alignment purposes of the hyperspectral images and to evaluate the degree of misalignment of acquired images, several of the most popular and frequently used algorithms were tested. One has to note that due to the very nature of the hyperspectral imaging process, the alignment of the images is similar to alignment of different layers of two RGB images. Comparing the performance of the alignment algorithms, the Scale-invariant feature transform (SIFT) resulted in the lowest registration error as shown in Figure 2.46D.

### 2.4.3 Watermark image extraction

After reviewing some aspects of the digital image acquisition and the image preprocessing, in the next sections some of the most applicable methods for extraction and processing of watermarks are presented. Some of the popular enhancement techniques are reviewed for completeness and to emphasize that they are necessary steps in order to generate a good (or sufficiently acceptable) quality image. Usually, preprocessing includes histogram equalization enhancement for adjusting image contrast and brightness. Histogram thresholding, including double thresholding, is an example of using histograms for image enhancement [133, 134]. It



is frequently used due to the natural differences in the pixels' intensity for the paper (background), the printed content, and the watermark. Image subtraction is used in combination with noise reduction and filtering in spatial and frequency domains. Algorithms for detection and extraction of chain lines in the image are based on the Hough or Radon transforms and edge detection methods, including Sobel, Prewitt, Roberts, Laplacian of Gaussian, and Canny [130 - 132]. The laid lines in backlight images appear as well-defined peaks in frequency domain due to their regular spacing and can be easily filter out.

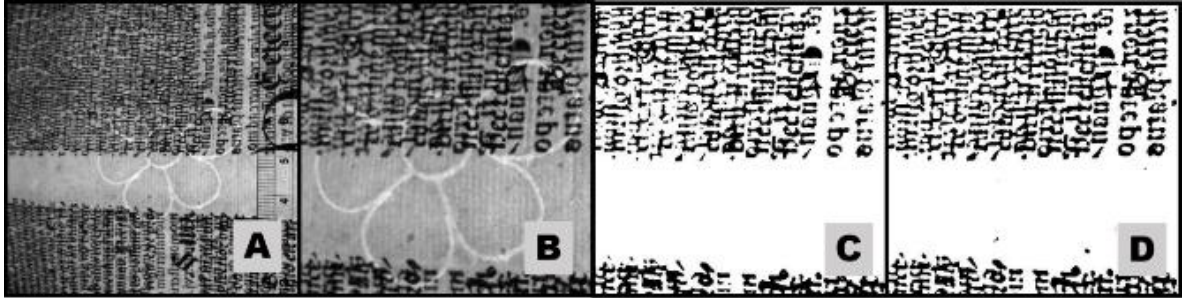
## **2.5 Watermark image preprocessing**

The digitized images of watermarks contain many unwanted interfering components such as text, chain and laid lines. The preprocessing of the watermarks images removes or at least decreases the interferences to produce an image of the watermark only. The steps typically involve thresholding to separate the background (paper surface with embedded watermark) and foreground (text and drawings). The task is complicated when an original *recto* image contains *verso* text. There are many techniques for layers separation and to determine an appropriate threshold value for the process of transforming a gray-scale image into a binary image for the object of interest. The techniques can be grouped into two broad categories: global thresholding and local thresholding. Global methods like that of Otsu [133] try to find a single threshold value for the whole document. Global methods are very fast and global thresholding statistical algorithms have been used for many years producing good results [134 – 139]. The global thresholding is not applicable to images of complex documents and with degradation due to uneven illumination and background interference. To overcome these complexities, local thresholding techniques have been proposed. As the name suggests, local

thresholding compares each pixel according to the grayscale intensity of the neighboring pixels. The techniques described and used by Bernsen [135], Chow and Kaneko [140], Niblack [141], Mardia and Hainsworth [142], Taxt [143], Yanowitz and Bruckstein [144], Eikvil [145], Sauvola & Pietikainen [148] and Singh et al. [149, 150] belong to the local thresholding category. The combination of global and local thresholding used by O’Gorman [146] and Liu [147] may be grouped as hybrid techniques. Reference [151] describes the performance results from a 2009 Document Image Binarization Contest of 35 research teams with 46 different algorithms which are also applicable to watermark images.

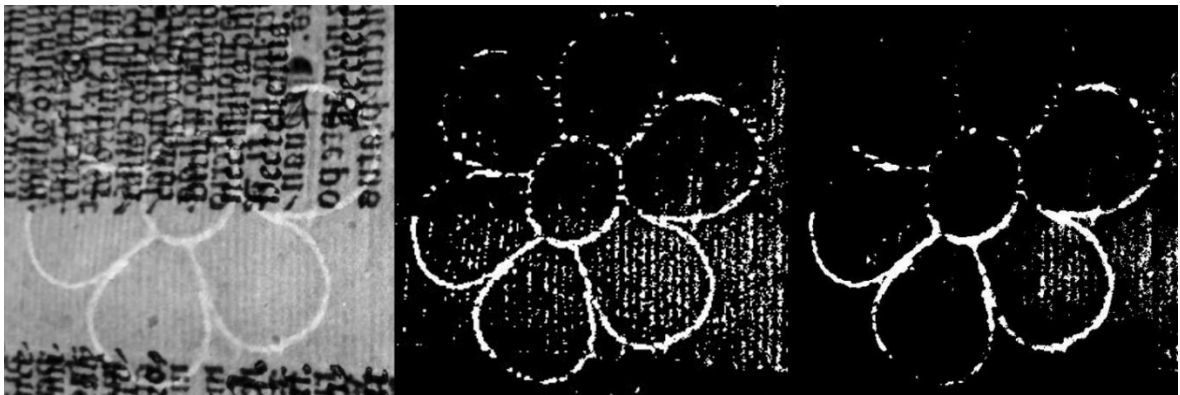
The use of preprocessing and binarization is illustrated with example images acquired from *Summa Theologica*. The four-volume set contains many pages with watermarks. The four volumes have been digitized using standard RGB reflected imaging and also with hyperspectral imaging.

Figure 2.47 presents the results of image preprocessing by applying global and local thresholding techniques to a recto image from *Summa Thologica* containing a very popular watermark of six-petal flower. The blue layer gray-scale of an RGB image was used (2.47A) with the region of interest (ROI) displayed in (2.47B). The background laid lines were extracted with multilevel histogram thresholding of the pixel intensities above the dark pixels of the printed text (black) and below the watermark pixels which have the highest intensity (white). The application of global thresholding (2.47C) and local thresholding (2.47D) “lifts” the printed text layer. Local thresholding performs better in classifying the character pixels.



**Figure 2.47: Blue layer of a *Summa Theologica* RGB recto image (A) and ROI (B). Global thresholding (C) performs worse than Local thresholding (D).**

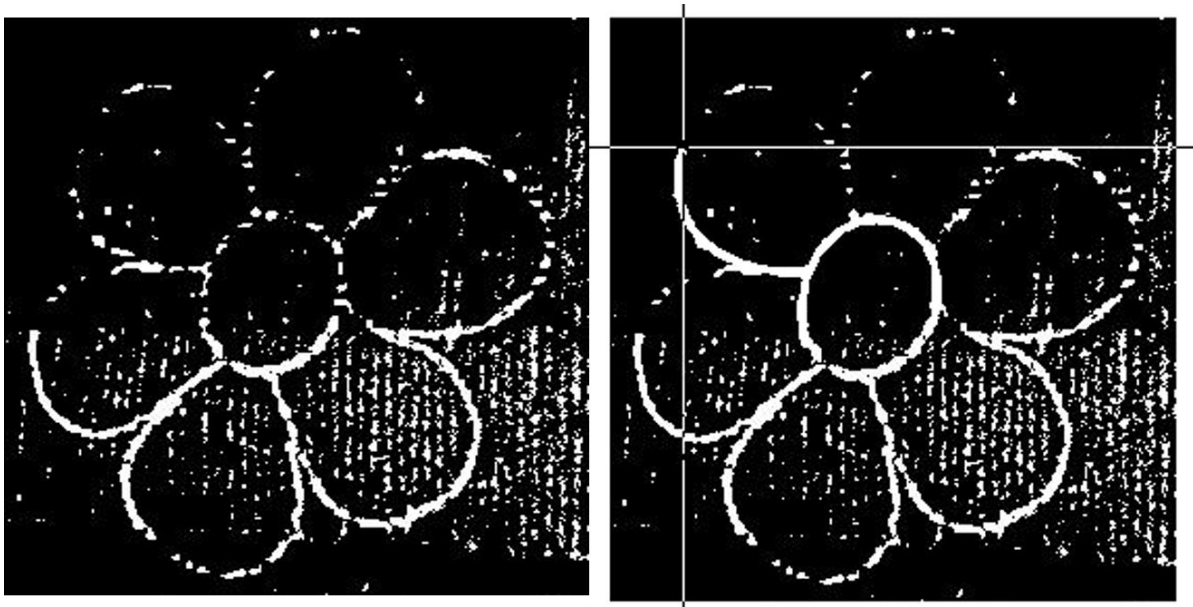
Figure 2.48 displays the original watermark image (left) and the watermark after the text was removed at different threshold levels (center and right).



**Figure 2.48: The original watermark image (left) and the watermark after the text was removed at two different threshold levels (center and right).**

The most commonly used processes in this watermarks related work are morphological operators. A structuring element with selected shape (square, disc, line, diamond, etc.) is used in combination with an image to perform sequential operators (e.g., union, intersection, difference). Applying morphological operations (e.g. dilation, erosion, opening, closing, reconstruction, and top-hat) modifies the image depending on the structuring element. The goal is to remove irrelevant data and simplify the image to extract the features of interest. There are many available methods [152, 153]. The ultimate goal is to produce a binary image of the watermark free of background interference.

Sometimes, morphological operators cannot restore completely the lines and curve segments of a watermark due to the limitations of the image quality. Figure 2.49 illustrates the interactive process of watermark (contour) restoration with high degree polynomial curves. In this particular example, a Bezier curve fitting is performed with the user selecting the number of points and clicking on a series of points of the petal contour. Bézier curve is a parametric curve frequently used in computer graphics to model smooth curves. Specifying the control points, the curve is completely contained in the convex hull of its points and can be graphically displayed and used to manipulate the curve intuitively. The curve is drawn directly on the original binary image. In Figure 2.49, the central circle has been restored and the cursor indicates the position of the currently selected point.



**Figure 2.49: Interactive digital restoration with high degree polynomial curves.**

The Bézier curves were developed in 1959 by the French mathematician Paul de Casteljaou for the design of Citroën automobiles, but became popular after the extensive use from Pierre Bezier to model the curves of Renault. Figure 2.49 is presented to illustrate the feasibility to interactively restore the lines of a watermark which is partially retrieved because of

interference with overlapping text. The image is processed in Matlab. Photoshop and some other image editors provide similar functionality. For example image restoration in Photoshop includes Bezier curves. Implementation with Java, Python or other open source language offers a free of charge ability for restoration. Other mathematical models could be used as well.

## 2.6 Watermark image binarization and skeletonization

In general, to generate a clear image of the watermarks only, the goal is to convert the original images into a binary form. The process transforms a gray scale image (or a monochromatic sub band of the hyper-spectral cub) to a binary image which contain only two classes of pixels: white as background and black as foreground.

The application of the preprocessing and thresholding techniques is demonstrated next with another image captured from *Summa Theologica Tertia (recto pp. 82, Volume III)* using reflected light imaging. The results of image processing are shown in Figure 2.50. The blue layer of the RGB image is used and the region of interest is selected. After applying histogram equalization and modified local adaptive binarization, the extracted binary image is almost free of print and background interference.

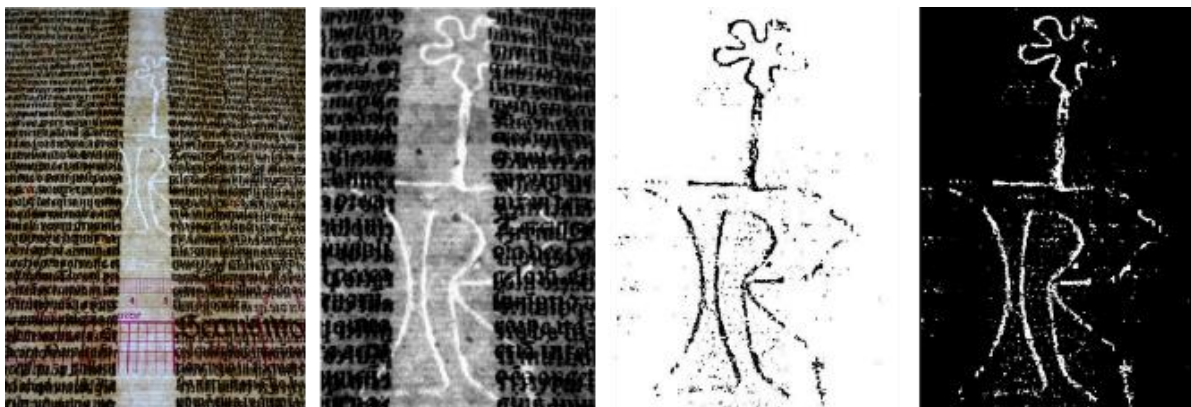


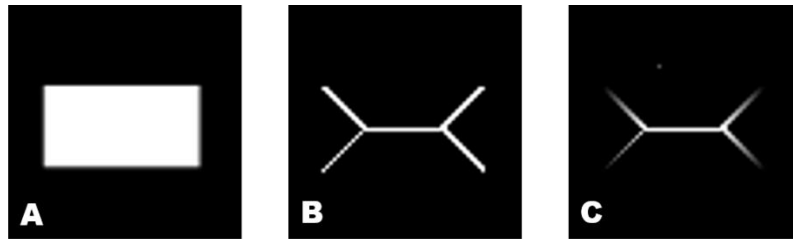
Figure 2.50: Image with watermark from Volume III of *Summa Theologica* (left) and the binary image of the extracted watermark.

The next step after digitization, segmentation, and binarization of the watermark images is to generate descriptors which can be used to represent and uniquely identify the watermarks. Currently, the methods for description, classification, and search are based on text, morphological criteria and hash tables, whole image comparison, and statistical features similarities [2, 12, 22, 154 - 155]. Some efforts with partial success were reviewed in Section 2.3 and presented in [95 – 98]. The search methods described in Subsection 2.3.3 illustrate the challenges and current deficiencies of these approaches.

The need to develop a descriptor that can faithfully represent the main distinct features of watermarks in compact format is obvious. A method for generation of a “template-like” descriptor (similar to applications in biometrics, for example) would result in a compact format representation with ability for fast and better comparison. In order to do that, the skeletonization technique is used next to further refine the watermark contours into a sequence of connected single pixels.

Skeletonization of a binary image is a process for reducing foreground regions to a skeletal residue that preserves the extent and connectivity of the original shape while removing most of the original foreground pixels. The thinning morphological operation produces a sort of skeleton. Thinning successively erodes away pixels from the outermost boundary (while preserving the end points of line segments) until no more pixel removal is possible. The alternative method is to first calculate the distance transform of the image. The skeleton then lies along the singularities in the distance transform. One disadvantage is the potential formation of curvature discontinuities. Another approach is to use the Medial axis transform (MAT) which produces similar (but not identical) results as the skeletonization and the two are often used with similar success (Figure 2.51). One advantage of MAT (not the pure skeleton)

is that it has the regenerative property that can be used to exactly reconstruct the original shape, if necessary.



**Figure 2.51: Binary image of a rectangle (A), and the result from the skeletonization (B) and Medial axis transform (C).**

Another way to think about the skeleton is as the loci of centers of bi-tangent circles that fit entirely within the foreground region being considered [116, 160]. The skeleton is useful because it provides a simple and compact representation of a shape that preserves many of the topological and size characteristics of the original shape. In respect to watermarks, the skeletonization is performed on contours (in general with more than one pixel width) and results in contours in which each pixel has two non-zero neighbors. At the cross-section of contours there are pixels with more than two neighbors. In the next chapter the need for skeletonization becomes more apparent when it is used in the process of generating a unique descriptor for the watermark.

The skeletonization process for a watermark (for a contour in general) can be defined as removing from its pattern as many pixels as possible without affecting the general shape of the pattern. The remaining pixels at the end should preserve the general shape and the pattern should still be recognized. The obtained pixels “skeleton” must have the following properties:

- Be as thin as possible;
- Be continuously connected (for a continuous shape);
- Be the shape’s centered representation.

Multiple algorithms have been developed to generate the skeleton of a shape. Using the Hilditch's algorithm, multiple passes are performed on a binary pattern and on each pass, the algorithm checks all the pixels and decide to change a pixel from black to white, if it satisfies the following four conditions:

1.  $2 \leq B(p1) \leq 6$
2.  $A(p1)=1$
3.  $p2.p4.p8=0$  OR  $A(p2) \neq 1$
4.  $p2.p4.p6=0$  OR  $A(p4) \neq 1$

If a pixel is changed, go to step 1. Else, stop.

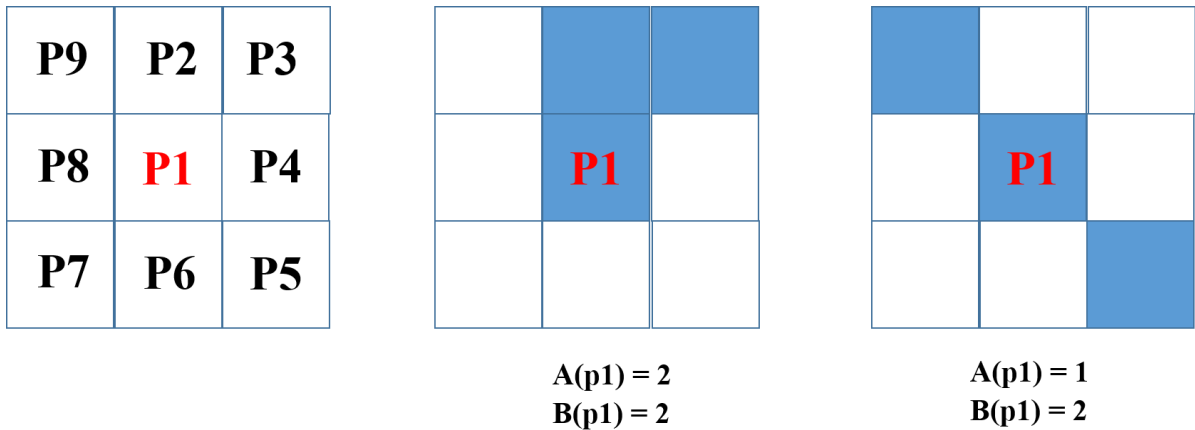
The meaning of the notation is given in Figure 2.52. Every non-boundary pixel **P1** has eight neighboring pixels **P2 – P9** (Figure 2.52, left). Please note that the indexed capital **Pi** indicates the location of the pixel and the small **pi** is its value of 0 or 1. The algorithm decides whether to remove p1 or keep it as part of the resulting skeleton. For this purpose, two functions **F<sub>a</sub>** and **F<sub>b</sub>** are defined for the **P1** neighboring pixels which are arranged in a clock-wise order to generate two values A and B:

**F<sub>b</sub> (p1)** = number of non-zero neighbors of p1

**F<sub>a</sub> (p1)** = number of 0, 1 patterns in the sequence p2,p3,p4,p5,p6,p7,p8,p9,p2

Examples for two possible patterns and the generated values A and B, using the functions **F<sub>a</sub>** and **F<sub>b</sub>**, are given in Figure 2.52.





**Figure 2.52: Examples for two possible patterns and the generated values A and B using the functions  $F_a$  and  $F_b$ .**

The Hilditch's algorithm is simple and easy to implement. It's a parallel-sequential process: all pixels are check at each pass and several passes are made until there is no change. Deficiency of the algorithm is that certain pattern will be erased completely and additional rules are needed to prevent for some of those possible pattern conditions.

For watermarks, the skeletonization was applied to reduce the binary contour(s) of the object to a single pixel wide representations. The original image is transformed to binary image foreground "1" and background "0". The algorithm works by making successive passes on the image. On each pass, border pixels of the contour are identified and removed on the condition that they do not break the connectivity of the corresponding contour. Figure 2.53 displays the original binary image of the watermark and the result of its skeletonization.

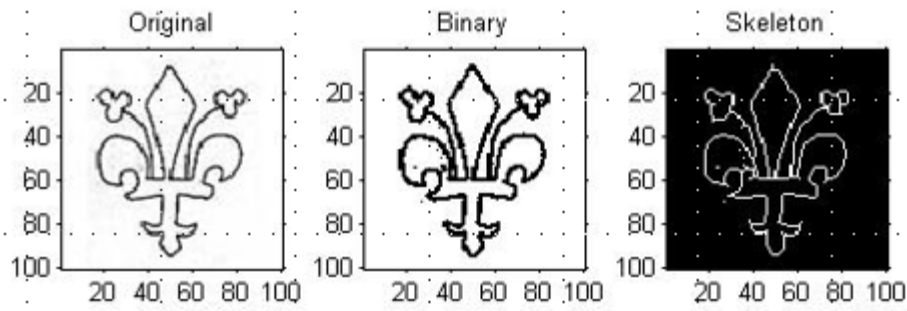


Figure 2.53. Lilie watermark original image, its binary image and its skeleton.

## 2.7 An alternative non-contact method for watermarks reproduction

The previous sections in this chapter demonstrated the steps involved in watermark image processing. Depending on the quality of the captured image, different techniques need to be used. The quality of the image depends on the state of the original document and the reproduction technology used.

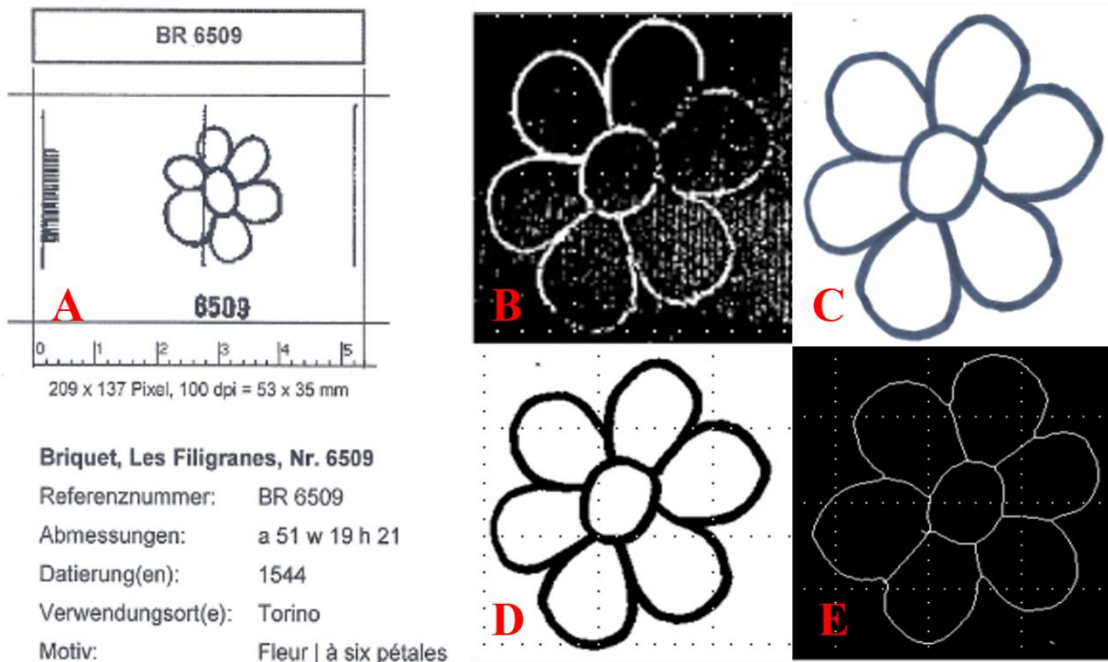
Non-contact capturing of the watermark image is always preferable in order to prevent any damages to the original document. In some cases, this is the only allowed method. Digital imaging, monochromatic, RGB, or hyperspectral is a very good methodology in this regard. However, even with *recto* and *verso* images, digital signal processing may not be able to successfully segment and extract a quality watermark. One possible solution is presented next. It consists of using an image of a document of interest. If a hyperspectral imaging is used, the optimal image is selected and printed. One common method for selecting the optimal band is by using pixel spectrogram comparison. Visually, the spectrograms of a pixel from the watermark and a background pixel are displayed and the wavelength with maximum difference is selected. The process could be automated with evaluation on multiple watermark pixels.

Next, a light screen is used to backlight the printed copy which can be easily traced to make a replica of the watermark as shown in Figure xx.



**Figure 2.54: LED backlight with the preprocessed image of *Fleur a six petals* watermarks (left). The tracing is performed with black marker (right).**

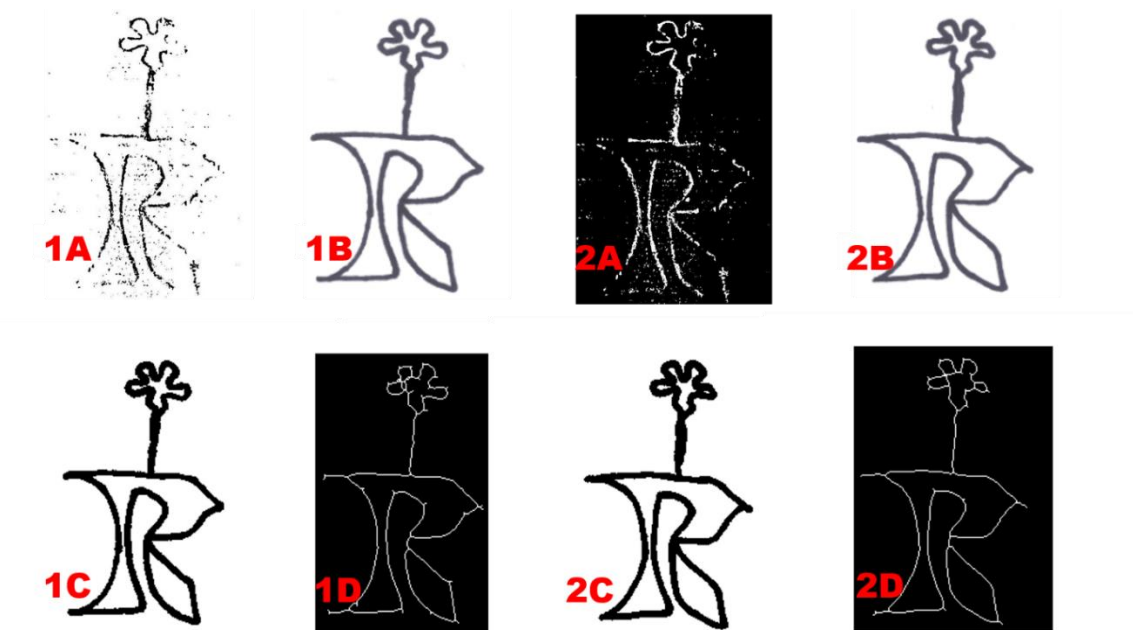
The tracing is done without direct contact with the original material – very valuable option for many document in fragile condition. The process could be repeated until a satisfactory result is achieved. Figures 2.55 and 2.56 display examples of watermark from *Summa Theologica* reproduced with this method.



**Figure 2.55: Briquet's *Fleur a six petals* watermark (A). A six-petal flower image from *Summa Theologica* after processing (B). The same watermark traced from the image (C). The watermark binary image (D) and its skeleton image (E).**

As can be seen, the process results in a very well-drawn watermark which are reproduced without a contact with the original document. In Figure 2.55, *Fleur a six petals* watermark BR 6509 from the Briquet collection is displayed on the left. In the center is the binary image of the same six-petals flower from *Summa Theologica* shown in the process of digital restoration with Bezier curves. On the right is the result of tracing performed with backlighting and a preprocessed image of the same watermark from *Summa Theologica*.

Figure 2.56 shows the results from the same reproduction method for another watermark from *Summa Theologica*, previously displayed in Figure 2.50. After the image was preprocessed, tracing was performed on the binary image and on the inversed binary image. Tracing the inversed image is for demonstration purposes. In some cases, the inversed binary image has the advantage of extensively blocking the back light, making the tracing process more accurate.



**Figure 2.56. The preprocessed *Letter R* watermark image (1A) and its inverse (2A). The two images after trace reproduction (1B & 2B), binarization (1C & 2C), and after skeletonization (1D & 2D).**

Figure 2.56 does not show any obvious differences between the use of the preprocessed original image and the tracing over the inversed image. Please note that the tracing of the six-petal flower displayed in Figure 2.55 was performed with a ticker marker. The tracing of the letter R watermark was performed with a *Sharpie* type pen. The quality of the paper and the markers used for tracing are important to prevent spreading of the ink and making the edges of the watermark fuzzy. The skeletonization process removes the effect of the ink spreading.

One practical advice is to use an image size that can be easily traced with uninterrupted continuous motion over a long portions of the watermarks. Tracing also could be performed with a digitizing panel and an active pen with a direct image overlay. It would be necessary to use a surface which responds to the stylus pen only and not the user hand contact. There are many available devices which can be used. In general, they are more expensive that the LED Light Pad (\$37) used for this work.

## **2.8 Conclusions**

The literature review on watermarks research and the selection of references in this chapter clearly indicate:

- the importance of digitization of historic documents;
- the amount and dimensions of the work already done in this area;
- the need to use better tools and techniques for imaging and image processing;
- the unresolved issue with an efficient description and retrieval of watermarks.

Some of the methods for watermark image processing were referenced and demonstrated by examples. The option for interactive restoration with segments drawn over the captured images was illustrated with good quality results. In addition, a new non-contact method for

watermark reproduction by tracing images of the original document was introduced with several examples.

The chapter introduced the capability of hyperspectral imaging for forensic analysis of historic documents. Applied with available software algorithms for dimensionality reduction, HSI becomes a very powerful tool for paper and paper watermarks research. After introducing the possibilities with the use of hyper-spectral imaging, in the next chapter the emphasis is on a new method for watermark description with the goal to allow an automated search, matching, and retrieval.

## CHAPTER 3

### AUTOMATIC IDENTIFICATION OF PAPER WATERMARKS

#### **3.1 Introduction**

Watermarks are the artefacts embedded into the paper during its production and caused by the intentionally added wire-made shape symbols, figures, or letters. Even without a close examination, it is easy to notice that most of the watermarks resemble contours, curvy-linear patterns, shapes, and simple or more complex figures. In this regard, any watermark image could be interpreted as a contour or already extracted object boundary. All methods used for description and representation of contour and boundaries are applicable to watermarks. It is highly desirable for a descriptor to be invariant to rotation, translation, and scale, easy to generate, to be compact and easy to compare, re-generative, and others [116, 160]. Additional criteria for descriptor selections are:

- Information preserving methods;
- Scalar transforms (compression) vs. space domain transforms (another image);
- Deterministic vs. statistical and probability methods.

As 2D data structures watermarks can be analyzed using shape factors and dimensions, shape moments, Contour slope sequence (CSS) and Fourier analysis of CSS, Radial function, Chain code, and combinations with additional techniques. More details for available shape description methods are presented next.

#### **3.2 Shape descriptors**

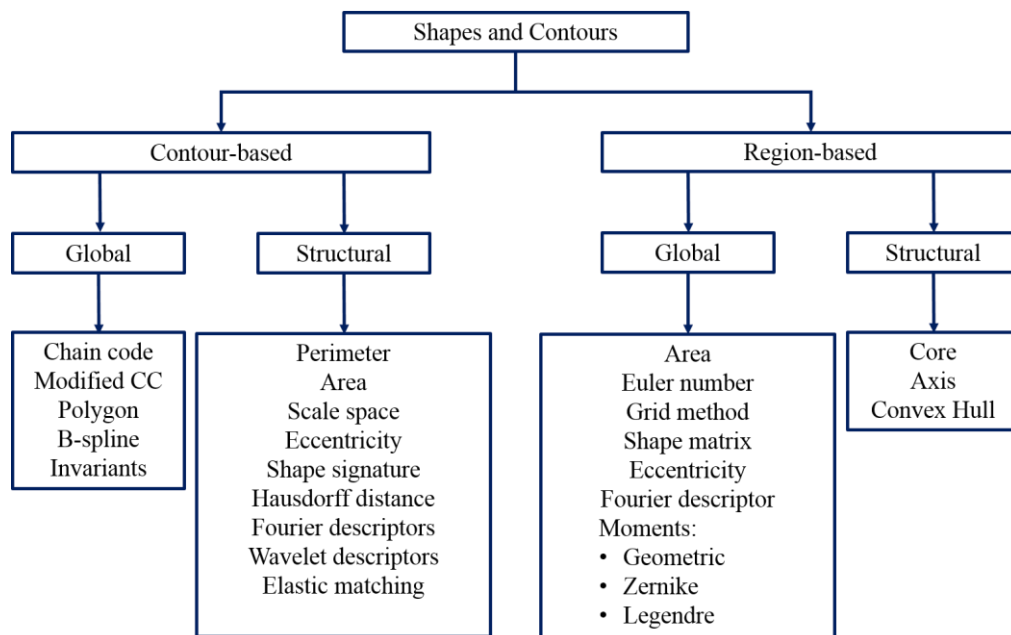
The generation of shape descriptors have been studied extensively and are well published and used [161-163]. Overall, shape descriptors fall into several categories: sequential boundary encoding, geometric representation, polygonal approximation (Guzman polygons), shape

moments, frequency domain descriptor (Fourier, wavelets, Curvelets, Gabor), and many others (Zernike moments, B-splines, etc.) [164, 165].

A good shape/contour descriptor should have the following properties:

- Good fidelity to the initial shape/contour;
- Good discrimination between different shapes and contours;
- Good behavior with shape/contour recognition operations :
  - Invariance with translation.
  - Invariance with rotation.
  - Invariance with scaling.

In this section, we provide a condensed review of available methods for contours (shapes) description for completeness and to justify our selection of the chain code as a descriptor for the watermarks. Figure 3.1 displays a block-diagram with a summary of the shape and contour description methods. Many of the methods have more than one implementation with enhancements designed for better and/or faster performance. An extensive review of shape representation and description techniques is given by Zhang and Lu in [166].



**Figure 3.1: Description methods for shapes and contours.**



**Simple contour metrics:** The simplest form of descriptor is the length of the boundary itself. Length of the major axis, basic rectangle formed by the major and minor axis encloses the boundary and the ratio gives the eccentricity. Area and the perimeter give the compactness ( $\text{perimeter}^2 / \text{area}$ ). Obviously, the metric has very limited power but could be used as a prescreening criteria during the matching process.

**Statistical moments:** Mean, variance, and higher order moments.

**Signature:** A 1-dimensional representation of a 2-dimensional object (shape, contour).

**Polygonal Approximation:** One classical method for compact boundary presentation is to fit geometric primitives (e.g. line segments) to the boundary and store only the parameters of these primitives instead of discrete points. This is useful because it reduces the effects of discrete pixelization of the contour. When the contours are thick (multiple pixels), the “tunneling” method is used when the straight line segments represent the contour by staying within its thickness.

**Contour slope sequence:** The tangent angle of the slope at every contour point is calculated relative to the slope at the starting point.

**Contour tangent curves:** If the direction of the tangent vector at every boundary point is calculated, we generate  $\psi(s)$  curve code of the tangent  $\psi$  as a function of the arc length  $s$  around the contour. Selecting a zero-degree tangent as a reference point, the  $\psi(s)$  curve will start and end at zero. The  $\psi(s)$  code is inherently translational invariant and close to rotational invariance with accuracy limited only by the segmentation method. The  $\psi(s)$  code could become also scale invariant, if the length of the contour is normalized (with some loss of resolution).

Comparing successfully two  $\psi(s)$  curves requires rotation of the  $s$  axis and shift of the  $\psi$  axis so that the starting point is at zero.

**Curvature and Bending Energy:** The first derivative of the  $\psi(s)$  curve describes the curvature at every point on the contour. The curvature representation is rotationally invariant. If the squared curvature points are integrated along the entire contour, the calculated single descriptor is known as bending energy. Bending energy is less descriptive but provides an additional discriminator which can be used for comparison.

**Extreme curvature points:** The extremes points are located where the contour describing an object has exterior or interior corners. Known as key salient points they are very useful for analysis and comparison.

**R(s) Curves:** If we calculate the length of the distances from a selected contour point to all other points on the contour, we compose a presentation known as the  $R(s)$  curve. Using correlation, the  $R(s)$  curves can be compared invariant to contour rotation.

**Radial function:** In a similar fashion to the  $R(s)$  curves, the selected reference point could be a different point inside the contour (usually the centroid). The polar shape representation is called the radial function.

**Cord Distribution:** Cord is any line connecting two points on the contour. There are many ways in which cords' lengths and orientations can be calculated and compared. The histogram of the cords' lengths is invariant to rotation. It is not scale-invariant but it scales linearly with the size of the object which contour it describes. The "angles" histogram of the cords' angles is invariant to object size but it shifts with respect to contour's rotation.

**Grid descriptor:** The cells of a constructed grid are assigned the value of 1, if they are covered by the shape (or covered beyond a selected threshold), and 0, if they are outside the

shape (below the threshold) [171]. A shape number consisting of a binary sequence is created by scanning the grid in a left–right and top–bottom order, and this binary sequence is used as the shape descriptor to index the shape. For two shapes to be comparable using grid descriptors, several normalization processes have to be done to achieve scale, rotation, and translation invariance.

**Delaunay Triangulation Method:** This approach of shape representation is histogram based. The corner points of a shape are used as the feature points of the object and a Delaunay triangulation of these feature points is constructed [164]. Consequently, a feature point histogram is obtained by discretizing the angles produced by this triangulation into a set of bins by counting the number of times each discrete angle occurs in the triangulation. A selection criteria for the angles ranges is necessary and varies with the shape types. The difference between two objects is the Euclidean distance between their corresponding feature point histograms.

**Medial axis transform:** MAT is composed by all the points of the shape which have more than one closest boundary points (“prairie fire” formation concept). It has regenerative property.

**Convex Hull:** The convex hull could be described as a minimal convex shape entirely bounding the contour. Selected metrics from the convex hull and the contour (mutual point, areas and perimeters ratios, etc.) may be used for additional features.

**Fourier Descriptors:** The method obtains the object representation in the frequency domain as complex coefficients of the Fourier series expansion of the objects shape signature. The method starts by obtaining a feature vector of the object called shape signature. The features could be curvature based, radius based, or boundary coordinates based. Next step is a

discrete Fourier transform of the shape signature: the Fourier coefficients are obtained and used for shape representation and for shape similarity calculation. For example, when the sequence of boundary points (the contour points) are treated as points on the complex plane, translation changes only the zero coefficient. Rotation in the complex plane is a multiplication by  $e^{j\theta}$ . The change in object size is equivalent to simply multiplying  $x(k)$  and  $y(k)$  by some constant. Changing the start point in the spatial domain (in this case,  $k$ ) is a phase-shift in the transform [165].

**Zernike Moment Descriptor:** The construction and use of orthogonal moments of an arbitrary high order based on the theory of orthogonal polynomials. Zernike moments are complimentary to chain codes, which are introduced next [164, 165].

**Chain code:** Introduced by Freeman in 1961 [167] chain code description of contours has some very intricate advantages. By encoding relative, rather than absolute position of the contour, the representation is relatively simple and translation invariant. It is also regenerative – the contour can be restored from its chain code without loss of information. In this regard, the chain code is a very efficient lossless compression method. Because the starting point pixel position in the image is necessary, one may consider the chain code as a combination of the coordinates of the starting point and the sequence of codes. Chain codes are boundary-based and represent the features of the shape's outline. For comparison, Zernike moments are region-based and represent the features of the whole shape so the two could be used in tandem to represent shapes [168].

**Angular radial transform:** The shape and contours of objects in images provide a powerful visual clue for identification and recognition of the objects. In particular, in binary images of watermarks, the shape feature is the most important visual feature – in general, there

is no color or texture in the image. Angular radial transform (ART), which is the region-based shape descriptor of MPEG-7, has the desirable properties for representing shape information using a small number of features. Because it has orthogonal basis functions, ART can represent the shape of an object in an image with no redundancy or overlap of information between the coefficients. The method employs a two-dimensional complex transform defined on a unit disk. It is rotation invariant, and has robustness to the noise. Unfortunately, it is computationally complex and a fast implementation is not available (yet).

In [168], Hodge et al. continue to expand their work on a neural network-based shape matching algorithm that uses Johnson Counter codes coupled with chain codes. They demonstrate how the binary associative-memory neural network can index and match chain codes which elements are represented by Johnson codes. With previous reports presented at conferences, the authors focus on shape representation after edge detection and application of chain codes as a better alternative to content based image retrieval. The proposed approach was developed for shape matching in trademark retrieval system.

The work of Hodge et al. is very relevant to the studies presented in this thesis. First, in [168], which was published in 2009, the authors have made a review of many shape coding and matching techniques. They conclude that “the recall accuracy for chain code shape retrieval was just below Zernike moments and well above invariant moments and edge histograms.” Second important aspect is the formulation of characteristics of the shape matching descriptor. The authors have made important statements regarding descriptor’s invariance and similarity. “A shape representation for figurative image retrieval should be invariant to translation, rotation and scale and similar shapes should have similar representations. Ideally the similarity of the representation should decrease monotonically as

the similarity of the shapes decreases.” [168, Section 3]. Third important aspect in the work of Hodge et al. is the obvious difficulties for chain code comparison. They have used several additional steps to represent the chain code and a binary neural shape matcher that represents shapes using a fusion of chain codes and Johnson Counter codes to allow distance-based matching. The fusion produces binary vectors which are integrated with a binary Random Access Memory (RAM-based) neural network (AURA) to allow rapid, efficient and accurate indexing and matching of shapes in shape databases. They convert the numeric chain code to a binary Johnson code equivalent, concatenate all Johnson codes together to form a vector, negate all vector elements and concatenate the original and negated vector to form the vectors for training the Correlation Matrix Memories (CMMs) are the building blocks for AURA systems [168]. The choice to combine the original vector with its negated presentation is questionable with respect to unnecessary data redundancy. The system is trained and tested with synthetic chain codes of equal lengths of 100, 200, and 300. Java random number generator in 0 to 7 range. It is not clear, if any rules are imposed on the generator in order not to compose unrealistic sequences (for example, after moving East to go back West for the next pixel). The point to make is that the codes are of equal length in order to be compared.

The issue of equal chain code length is addressed by other researchers as well. Hodge et al. referred to the work of Lu [169], who introduced procedures for normalizing the C8 chain codes to ensure invariance to rotation and scaling. He has proposed orienting shapes along the principal (major) axis of the shape. With a minor axis perpendicular to the major axis, this forms a superimposed rectangle which may be subdivided into a grid of cells. To ensure scale normalization, the authors in [168] have fixed the perimeter-size of this grid so that it is equal for all shapes (for example, 4x4, 3x5, 2x6 or 1x7 grid cells all have equivalent perimeter length

of 16). Superimposing the most appropriate of these cells onto the shapes and starting from the top right cell, the shape presentation is normalized for rotation and scale and represented with equal length.

Both the grid and the object sizes affect the binary number derived for an object. Therefore objects must be normalized for scale with a fixed length of standardized major axis. Subsequently, scaling normalization is performed along the major and the minor axes proportionally. The minor axis is perpendicular to the major axis and of such length that a rectangle with sides of major axis and minor axis defines the minimum bounding rectangle of the object.

Rotation changes the spatial relationships between the binary grid cells and the objects boundaries. This leads to different binary number representations for the same object. Therefore objects must be normalized for rotation. This is achieved by obtaining the major axis of the shape and subsequently rotating the shape to make the major axis parallel to the x axis. The rotation and scale invariance is very important and it will be revisited further to point out the deficiencies in the existing methods and propose a solution with a new method.

**Gabor wavelet, curvelet transform, edge-direction histogram and dendrograms** methods have been applied to shape and contour description and matching with different degree of success [116, 157-163]. In general, the description methods can be divided in two types of very different approaches for contour shape modeling: conventional and structural. Conventional approaches treat the boundary as a whole, and a feature vector derived from the entire boundary is used to describe the shape. The measure of shape similarity is usually the Euclidean distance between the feature vectors. Structural approaches use certain criterion to

divide the shape boundary into segments, known as primitives, which are used to construct a string or a tree. The measure of shape similarity is string matching or graph matching.

After reviewing the available contour descriptors and evaluating their performance on watermark images (including possible improvements), the chain code was selected for generation of unique template (signature) of watermarks. The next section presents the advantages of the chain codes: simplicity of generation and its regenerative property. It will be shown later that the regenerative property of the chain code allows a complete description of more complex images and nested contours by applying the process iteratively with consecutive removal of the described portions. This approach cannot be used with other shape and contour descriptors.

### **3.3 Chain codes as an image descriptor**

Chain codes describe shapes using sequences of numbers. They are simple and flexible. Before exploring further the use of chain codes for the description of watermark, the original Freeman idea is presented. In the 2D space of an image, every pixel has eight neighboring pixels (Figure 3.2B). If the pixels are numbered counterclockwise from 0 to 7, the relative position of each of the neighboring pixels can be represented with a single number. The 8-directional chain coding schema is shown on Figure 3.2B. In the literature, this is frequently referred as the classical Freeman configuration with eight neighboring pixels [167]. The selected labeling characters are actually irrelevant and one may choose any characters to map the pixels and directional connectivity. The orientation of labeling (clock-wise or counterclock-wise) and the position of the 0 are not strictly defined. The lack of standardization obviously results in non-compatibility.

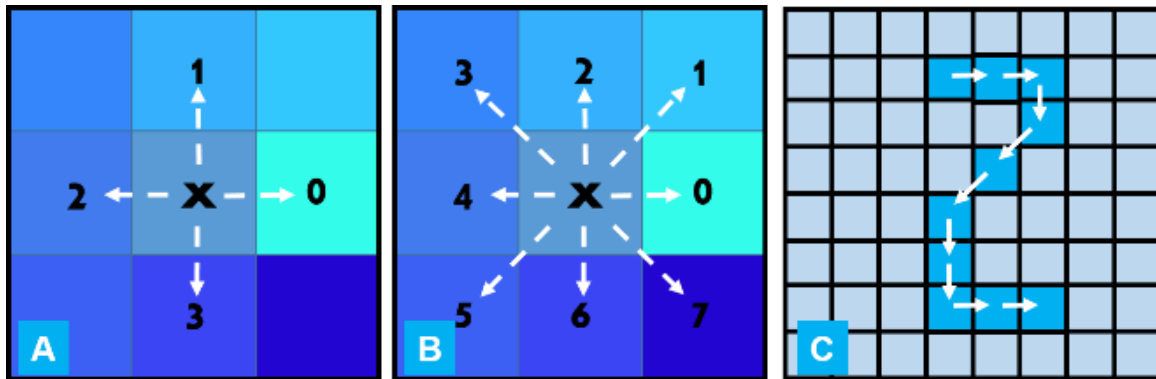


The methods for generating the classical chain codes (CC) were introduced and described in [167]. Frequently it is referred to as F8 or as Freeman 8-directional chain code. It is the most intuitive chain code. Introducing the 8-pixel connectivity, Freeman also specifies that the boundaries of digitalized shapes can be described by the horizontal and vertical neighboring pixels. If we restrict the relative description to only horizontal and vertical directions (Figure 3.2A), only four characters are necessary, which can be represented with two bits only. Historically, 4-directional CC (four-pixel connectivity) for pattern classification were introduced by Duda and Hart in 1973 [170]. With 4-directional CC we lose the ability to represent the diagonally positioned neighboring pixels. In addition, the A8 chain codes have been developed in which the angle for motion to the next pixel is coded. The result of representation is very similar: an 8-character alphabet is used and processing techniques are equally applicable to F8 and A8 chain codes [168]. It is important to note that A8 representation compared to the F8 representation has the advantage of being rotationally invariant while the F8 is not.

The 8-pixel connectivity chain code is used in this study for representation of paper watermarks. During the generation of the chain code, the algorithm starts from the top most left pixel and, moving counter-clockwise, evaluates the neighboring pixel position with value 1. Formally, the CC generation can be defined with the following notation:

$$CC(i, j) = \begin{cases} 0, & p(i, j + 1) = 1 \\ 1, & p(i - 1, j + 1) = 1 \\ 2, & p(i - 1, j) = 1 \\ 3, & p(i - 1, j - 1) = 1 \\ 4, & p(i, j - 1) = 1 \\ 5, & p(i + 1, j - 1) = 1 \\ 6, & p(i + 1, j) = 1 \\ 7, & p(i + 1, j + 1) = 1 \end{cases} \quad (3.1)$$

The CC generation is illustrated using the synthetic pixel representation of number 2 on an 8 x 8 pixels grid (Figure 3.2C). In this and future examples, we selected as a starting pixel to be the first pixel in the binary image scan from left-to-right and top-down. Following the simple rules, the chain code for binary image of #2 in Figure 7-C becomes  $\{(2, 4), 0, 0, 6, 5, 5, 6, 6, 0, 0\}$ . The (2, 4) pair captures the indexes of the start pixel, which are row 2 and column 4, and follows the Matlab convention of indexing the top left corner has index (1, 1). These indexes will be offset by -1 if the chain code is generated in different programming language (C++ or Python, for example) because the frame indexes start at (0, 0). From the starting pixel indexes and the chain code sequence, the coded image can be restored (regenerative coding). The original directionality assignment is also important (clock- or counter-clockwise). Unfortunately, not only the direction but also the location of the starting alphabet character has not been standardized.



**Figure 3.2: Chain code generation schema with four and eight neighboring pixel.**

Nunes et al. [171] encoded the boundary edges of the shape instead of the pixels and therefore the walking direction should be specified. They proposed differential chain code (DCC), the first chain code with only three elements. Namely, the DCC alphabet is R for right, L for left, and S for straight. In 1999 Briabiesca introduced a new, nowadays very popular chain code, named the vertex chain code (VCC) [172]. While developed to resolve some of issues

with classical chain code, it has somewhat restricted applications and RST-variance is not resolved. VCC determines the number of boundary pixels in the raster junctions. It has an alphabet with only three codes. The Three Orthogonal (3OT) chain code with 3 symbols only was introduced in 2005 [173]. Its codes are determined as follows:

- 0 - when the current coding direction is the same as the coding direction of its predecessor;
- 1 - when the current coding direction is equal to its first predecessor whose coding direction is different than the direction of its predecessor;
- 2 - for all other directions.

Some further developments of the VCC were introduced in [174]. In [175] the authors compare the classical (Freeman) CC with the VCC and describe a string-edit distance for shape boundaries comparison. They also suggest statistical approaches to solve the inherent rotational and scale variability of the chain codes. A partial solution to the rotational dependence of the CC is to encode not the relative direction of the next neighboring pixel, but the differences in the successive directions, generating a differential chain code (DCC), as proposed by Nunes et al. [171]. This can be computed by subtracting each element of the CC from the previous one and taking the result modulo  $n$ , where  $n$  is the connectivity used to generate the CC (4 or 8). This DCC resolves comparison of object rotations only in +/-90 degrees increments. An example in Table 3.1 illustrates the generation of CC and its DCC.

**Table 3.1: Example of CC, DCC, and shape numbers for a simple contour.**

<b>Classical Freeman CC</b>	<b>0 0 6 5 5 6 6 0 0</b>
<b>The difference CC</b>	<b>0 2 1 0 7 0 6 0 0</b>
<b>The shape number</b>	<b>0 0 0 2 1 0 7 0 6</b>

A possible solution to invariance with respect to the starting point is to rotate some portion of the DCC so that it represents the smallest value viewed as an L-digit integer. Such a

normalized DCC is called a shape number (Table 3.1). Shape numbers are very sensitive to rotation, scaling, and to discrete pixel artifacts. Researchers have shown that resampling the grid along some principal axis of the contour (at some angle and sampling rate different from the original grid) can potentially account for rotation, resizing, and minor pixel artifacts. This does not provide a robust solution with repetitive results.

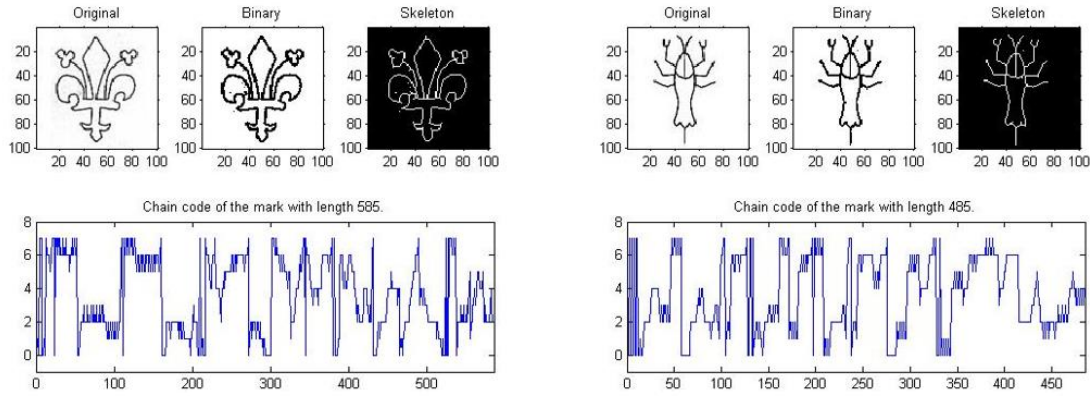
Chain codes have been applied as representation methods for rasterized shapes within various scientific and engineering disciplines [176–184]. The chain code can be regarded as a sequence of commands, which control the clock-wise movement of a boundary (contour) pixel walker. Introducing the 8-pixel connectivity, Freeman also specifies that the boundaries of digitalized shapes can be described by the horizontal and vertical neighboring pixels only but with less resolution.

All shape and contour descriptors reviewed above could be considered as a 1D function representing a 2D object boundaries and contours. CC, DCC, shape numbers,  $\psi(s)$  curve,  $R(s)$  curve, and the other descriptors could be used as a signature or feature vector representation. They have different advantages and shortcomings. In respect to potential signature presentation of digitalized paper watermark, we are most interested in generating unique signatures that are RST-invariant, easy to compute and compare. In the next section, new method is presented for constructing such descriptor.

### **3.4 Chain code generation for watermarks (from an image to a chain code)**

Next we illustrate the steps for generating RST-invariant CC for the Lilie watermark [69]. As presented in Figure 3.3, the original gray-scale images are transformed to binary images. The next step is tinning of the watermark to reduce the binary image to a neighboring sequence

of single pixels. For this purpose, we used skeletonization morphology operation [116, 160, and 185]. We use the classical F8 Freeman CC with a starting point the first pixel of the watermark detected with left-to-right and top-down scanning of the binary skeleton image.



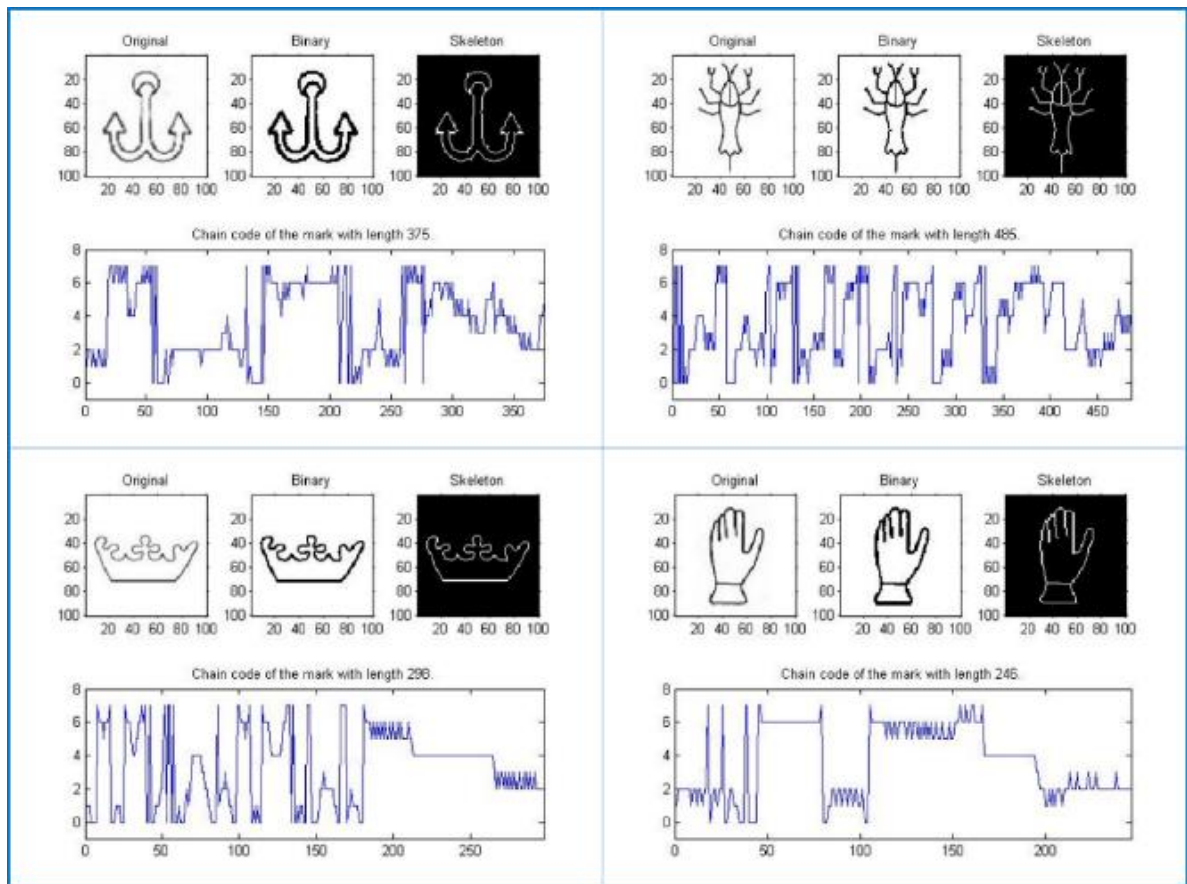
**Figure 3.3: CC generation for two watermarks after binary and skeleton transforming of the original images.**

Next, Figure 3.4 presents the results of the same process for four other watermark from [69]. Even at a smaller scale, the images demonstrate visually how different the generated chain codes are. The chain codes are displayed as graphs using the code number as a magnitude value. The watermark have different contour lengths and respectively different CC lengths of 375, 485, 298, and 246 for the *Anker*, *Vielfuber*, *Krone*, and *Hand* watermarks.

The definition of *filigram* is “a distinguishing mark impressed on paper during manufacture, visible when paper is held up to the light.” (<https://www.vocabulary.com/dictionary/nl/filigram>). *Filigram* was given as a synonym of the French *filigrane*. Simon Tatham, a mathematician and programmer, has introduced “filigrams” as “a type of pretty pictures generated by mathematical means” (<http://www.chiark.greenend.org.uk/~sgtatham/filigram/>).

“Filigram ID” or “Filigrane ID” could be good choices to use as term for the chain coded watermark. However, after considerations, I have decided the representation of a watermark with CC to be introduced with the new terms, as follows: “chain code filigram” (CC-filigram)

or the easier to use “C-gram”. The motivation for the terminology is further explained in Chapter 7.

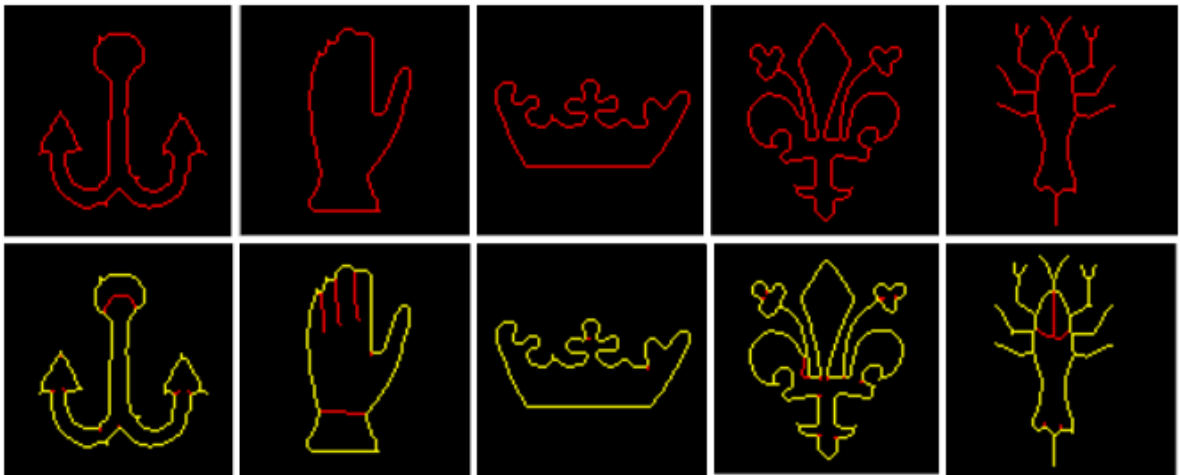


**Figure 3.4: The Chain codes of the four watermarks are very different.**

One has to note that at this point the introduction of the CC as a descriptor is based on the single iteration involving the outermost contour of the watermarks. The application of CC for complete watermark representation is introduced and described in detail in Chapter 7. The process is further illustrated in the next section, where the reverse transform is used to form back the original image from a generated CC.

### 3.5 The reverse transform (from chain code back to the original image)

The re-generative feature of the CC is illustrated on Figure 3.5. The re-generated watermark images are displayed on the first row in red color. They are superimposed on the original images using yellow color. The portions of the original watermark images, which are not represented with the first iteration of the chain code generation, are in red. The remaining “red” differences are the segment to be processed during the following consecutive iterations. The CC algorithm will process them in left-to-right, top-down order. Also, the CC generation algorithm is design to follow clockwise the outermost contour. That is why all differences occur internal to the contours of the watermarks.



**Figure 3.5: Watermarks restored from their CC (top row) and displayed with yellow color on the bottom row for comparison with the original watermarks. Differences are in red.**

For the relatively simple watermarks in Figure 3.5, the differences, resulting from the original images and their contours, do not appear to be a substantial portion of the images. The chain codes based on the contours would not cause a matching problem for this small set of different watermarks. However, describing the watermarks completely is important when comparing very similar watermarks. We address this problem with the formation of a

Composite feature vector (CFV), introduced in Chapter 7. Here, the emphasis is on the re-generative property of the chain code.

The re-generated image is an exact copy of the original image. This provides the ability to iteratively describe even the most complex watermark. The CC generation is followed with a re-generation of the described portion of the original image. This regenerated image is subtracted from the original and the next CC generation is performed on the remaining image. The impatient reader may refer to Chapter 7, which represents the iterative steps for complete description of complex images.

### **3.6 Chain code as a lossless compression**

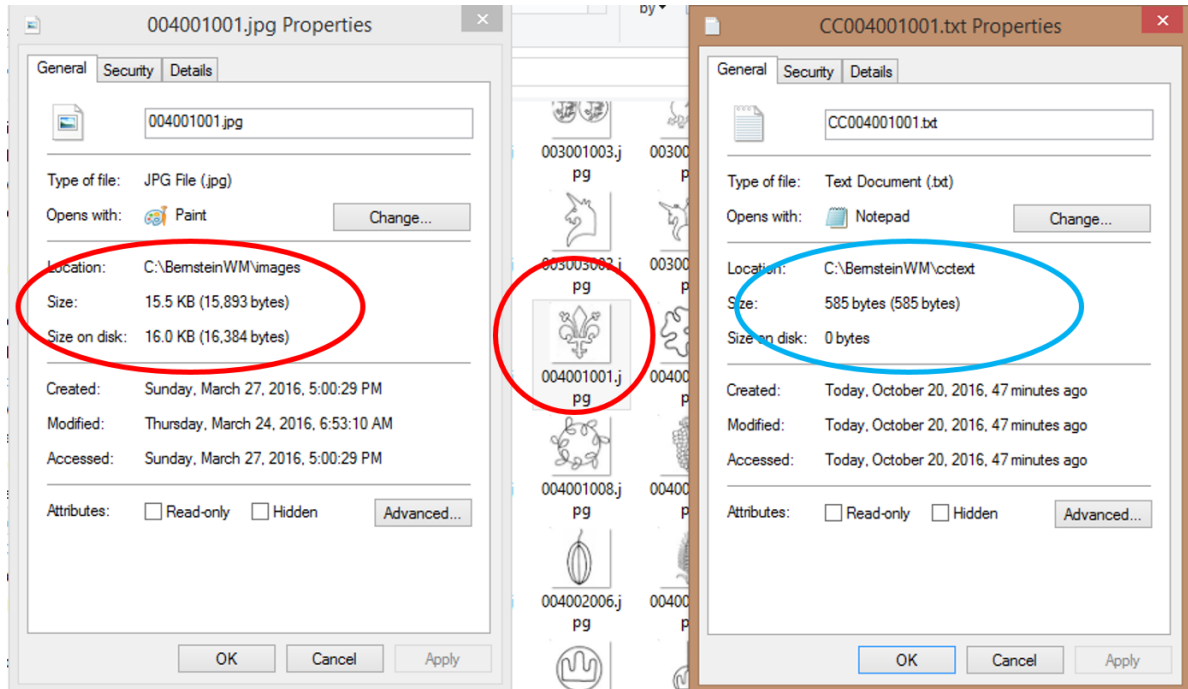
The pioneering work in the chain code compression has been done by Nunes et al. [171], who used Huffman codes on DCC [186] to achieve a near-lossless compression. Liu and Zalik [187] also used Huffman-based compression on their directional difference chain code (DDCC). Liu et al. [188] introduced three methods for vertex CC (VCC) with symbol combinations. Sanchez-Cruz et al. [189] introduced a modified 3OT (M\_3OT) chain code, by introducing additional symbols in order to encode frequent combinations of 3OT chain code symbols 0 and 1. Sanchez-Cruz [190] have proposed a newer compression method for F8 by finding repetitive substrings within the chain code. Liu et al. [191] have proposed a quasi-lossless chain code compression method, which replaces less frequent angular differences in DDCC with pairs of more frequent ones. They have also introduced a run-length encoding (RLE) scheme in order to further compress repetitive runs of 0\_ angular differences. Recently, Zalik and Lukac have presented in [192] a new lossless chain code compression that does not rely on statistical assumptions (i.e. Huffman coding tables), but rather use a move-to-front transform (MTFT) to reduce the information entropy and then encode the transformed symbols



with adaptive run-length encoding (ARLE). According to the authors, currently this method provided the most efficient compression of VCC, 3OT, and normalized angle difference (NAD).

In [222], Zalik et al. have presented a “universal” chain code compression method. The authors have described their universal method based on three-mode algorithm. First, the chain codes are converted into a binary stream, independent on the input chain code. Different binarization schemes are used for different chain codes, leading to universal input for further compression. Then, the compression is done using three modes: RLE0 compresses the runs of the 0-bits; LZ77 handles the repetitions within the bit stream; and COPY is an escape mode used, when the other two methods are unsuccessful. The authors conclude that the Freeman chain code in eight directions (F8) is more compressible than Freeman chain code in four directions (F4), the Vertex chain code (VCC), or even the Three Orthogonal chain code (3OT). The most compressible chain code is the normalized angle difference (NAD) chain code but it is sensitive to roughness of the geometric shape.

After the watermark images are transformed to binary images and the skeletons are generated, the number of foreground pixels are substantially smaller compared to the size of the image. Furthermore, for simple contours watermarks, this number of foreground pixels is close to the length of the chain code for the contour. In the example cases presented previously, the original images are 100 x 100 pixels in JPG format resulting in sizes from 11KB to 19KB. The computed chain codes recorded as a text files are less than 1KB in size. Figure 3.6 shows the properties for the original JPG image (circled in red) and the properties of the chain code TXT file (circled in blue).



**Figure 3.6: Comparing the size of the original watermark jpg image and its chain code.**

The result for this average size JPG file is an additional compression of about 27. As discussed previously, the compression is lossless because of the re-generative property of the chain code. With the development of the RST-invariant CC, the restored images are identical to the original images.

Another advantage of the chain coding can be seen by pointing out that it needs only 8 characters. With other words, 8-pixel CC needs an 8-character alphabet. When these characters are chosen to be the decimal digits 0 - 7, they can be represented only with three bits. Figure 3.7 displays the CC for the watermark image file 004001001.jpg from the Bernstein dataset. The length of the CC is 585. These 585 decimal values can be converted to binary with a fix length of 3 bits. The 3-bit groups can be concatenated to form a sequence of any desired length. Additional compression of the original JPEG file can be achieved with decimal to binary conversion of the CC, followed by concatenation of the groups of three bits to form a sequences of desired length. The sequences can be represented with a higher basis for even greater

compact representation and storage. Tables 3.2 shows the original chain code for Bernstein #004001001 watermark with length 585. The chain code has indexes  $i = 1, 2, 3, \dots, 584, 585$ .

**Table 3.2: The original chain code for Bernstein #004001001 watermark (585 bytes).**

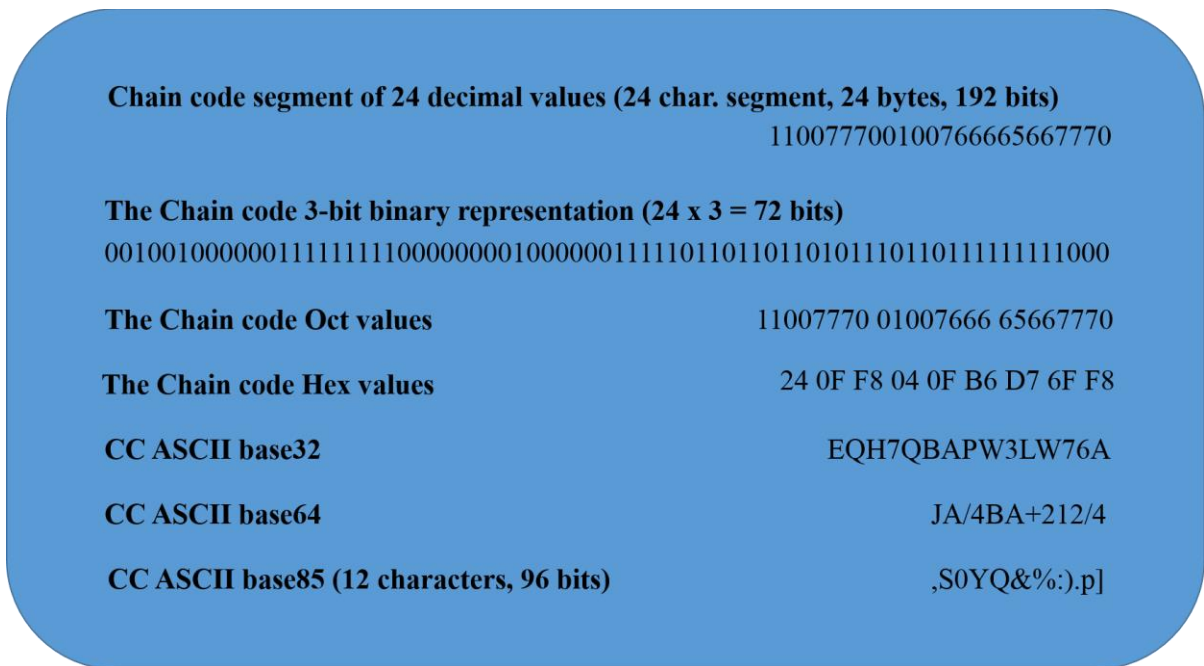
<b>index</b>	1	2	3	4	5	6	7	8	.	i-1	i	i+1	.	585
<b>value</b>	1	1	0	0	7	7	7	0	.	.	.	.	.	2

The chain code with values after decimal to 3-bit binary conversion ( $i = 1, 2, 3, \dots, 584, 585$ ) is shown in Table 3.3.

**Table 3.3: The original chain code represented in binary format (~220 bytes).**

<b>index</b>	1	2	3	4	5	6	7	8	.	i-1	i	i+1	.	585
<b>value</b>	001	001	000	000	111	111	111	000	.	.	.	.	.	010

The chain code with an index value from 1 to 585 has a length of  $585 \times 3 = 1755$  bits in binary representation or approximately 220 bytes. If the standard ASCII representation is used to convert sequences of eight bits to characters, some 8-bit groups will be non-printable characters. One possible solution is to use higher than eight base representation. Figure 3.7 demonstrates the decrease of representation length with the increase of representation base.



**Figure 3.7. Comparing the size of the watermark's CC with different representations.**

It is well known that increasing the encoding base from 16 (hex) to 32 and 64 increases the efficiency of representation. Unfortunately a Base128 encoder would not have sufficient printable ASCII characters: only 94 are available, excluding the whitespace. While creating a base 94 encoder theoretically would be more efficient than the lower bases, it turns out that base 85 is just as efficient with some implementation advantages which are outside the scope of this thesis.

The character patterns in Figure 3.7 indicate that ASCII base 32 base encodes the original 72 bits (9 bytes) as 15 characters. ASCII base 64 encodes the original 72 bits (9 bytes) as 12 characters. This increases the data size in bits by about 33%. For this particular example, the ASCII base 85 encodes the original 72 bits with 12 characters as well. However, for longer lengths, ASCII base 85 represents 4 bytes as 5 characters, which increases data size by just 20%. For example, the ASCII base 85 representation of **0010001000000111111111000** (which is the concatenated string of the first 8 chain code values of the Bernstein #004001001 watermark) is **+p\#G**. Furthermore, ASCII base 85 representation, is using an 85 character alphabet which are all printable.

Typically, ASCII base 85 is using 32 bit groups. Figure 3.7, presents a 24 characters segment which is the initial part of the Cain code for the Bernstein watermark #004001001 in decimal values. The octal and hexadecimal coded values, and the ASCII conversions with base 32, base 64 and base 85 representations. The application of ASCII base 85 representation adopted for watermark chain codes in this work uses 24 bit groups because of the 3-bit representation of the 0 to 7 numeric decimal values. Simple consideration is the simultaneous divisibility by 3 and also by 8 (common denominator). To summarize, a 24-bit group represents the F8 CC values with a three bytes. The next possible length divisible by 3 and 8 is 48 but the

gain efficiency is minimal and the additional considerations described next made the selection of 24-bit groups for more efficient (compressed) representation. The represented study is only an exercise to extend the lossless data compression ability of the CC.

If the CC values of the Bernstein watermark #004001001 are presented with 3-bit binary values and bits are grouped in 24, there will be  $585 \times 3 = 1755$  bits and  $1755 / 24 = 73.125$  groups. The preference is to have an integer number of groups. The natural thing is to append zeroes to the CC. which equals to adding 21 zeros at the end. So  $1755 + 21 = 1776$  for a total of 74 groups of 24 bits. At the beginning of the CC, another group of 24 bits is added with the first 12 bits giving the x pixel coordinate of the original image in 3-bit binary number presentation. The second 12 bits represent the y-coordinate of the CC's start pixel. With 12 bits the size of the images could be  $4096 \times 4096$ . A second (overhead) group of 24 bits is added after the very first one to indicate the number of zeroes used for padding. The result is  $1 + 1 + 74 = 76$  groups of 24 to represent the watermark with length 585 bits. The actual overhead is  $24 + 21 = 45$  bits because the first group with the start pixel coordinates is essential. Table 3.3 presents the grouping of the binary bits for the first 8 CC values: 11007770. This forms four groups of six bits each. When they are concatenated, group of 24 bits is formed ready to be represented with ASCII base 85.

**Table 3.4: The original chain code forming the first 24-bit group.**

<b>index</b>	1-2	3-4	5-6	7-8	.	k-1	k	k+1	.	.
<b>value</b>	001001	000000	111111	111000	.	.	.	.	.	.

In general, ASCII base 85 doubles the compression achieved with the CC representation. The CC length depends on the complexity of the watermark. For the Bernstein watermark #004001001, the ASCII base 85 results in 280 bytes representation of the 585 bytes (numeric decimal characters) CC. Overall, the CC compression ratio of the 15893 bytes jpeg image file

is reduced to 280 bytes or  $15893 / 280 = 56.76$  or more than 50 times compression. The compression capability is demonstrated as a potentially very useful characteristic of the CC which could be applied when the size of the stored data is the primary concern. In addition, error correction could be included, if desirable or necessary.

Figure 3.8 displays the CC for Bernstein watermark #004001001 expressed as decimal numeric characters and its ASCII base 85 representation. Please refer to Appendix C for details of the example. One may notice that the representation with ASCII base 64 results in many repeatable characters, a factor that is typically used for further compressed representation.

One advantage of the proposed compression is the simplicity of implementation compared to more complex methods. One possible disadvantage of additional representation step for reducing the size of the CC is the need for reversed conversion in order to take advantage of the developed CC distance measure which is based on the numeric decimal characters representation.

Chain code represented with decimal characters: 585 bytes.

```
110077700100766665667770776776766676676666566766766700012222322223222232232232
22322323222211212112121121122667077677676777676766565656566656566656656666
56670000012222212222222212212121111111011232222100770110007666654456667444443
2344545555555565656656666666656666701221222212221211101100000007777667666656
6565554455122223444456556677076666544434444434444566666666566670070000115666
545444566666554432332222344443322326700000122122222223444445444444434401101
12222233444555666773444344333322222222212111010007007077776766701222232232
22323451223343345554444522322222343222
```

Chain code represented with ASCII base 85: 280 bytes.

```
,S0YQ&%:).q#>iXg[4^gmIpEr!D+W$9dH0%-B`("dR]U".&'QfD9)PPp]#[8]'HNN?+&:"[G
,K+!!#ng'Tu$5;CJ82,'kW4,XD8?!;gVD]5JsWV^7Y`D7]&@gMOCND7D#/OsqT3+HQR'&-(D
O[G'q`P1#LbaX,O*q"no=N_(IY9=ESIG2*Q-"PpJO8TrimA8/E6:.ZBrXhXoNMFf(WP"B2I8
E9o*;Fr4rQJK%<8E=WpP"?p]7$7fXJ]IfZ[fn%G9dJ2PUtfilA<Nb:8N>,>
```

**Figure 3.8: CC for Bernstein watermark #004001001 expressed as 585 decimal numeric characters and its ASCII base 85 representation with 280 characters.**

As recently shown by Sanchez-Cruz and Lopez-Valdez [24], different type chain codes can be directly converted from one to another. The use of different chain codes may add additional variability without sufficient gain to justify it. However, if needed, the conversion feature can be efficiently included in the proposed new algorithm for watermark representation.

### **3.7 Limitations of chain codes**

Due to the method of their generation, chain codes (CC) for the same contour are different at different image orientation and scale. Rotation variance will generate different sequence. Different size of the image will result in a different CC sequence with a different length. The variability with scale and rotation has been the main limiting factors for the broader application of the chain codes. To address the RST-variance of the CC, the authors of [224] propose minimized sum statistical direction code with a similarity measure based on its entropy but the presented results are very limited and not conclusive. In [225], chain code coherence vector and chain code distribution vector are presented and evaluated with proposed similarity metrics. In addition to statistical features CC distribution is considered to make the method invariant to translation and rotation and proportional to scale.

### **3.8 Conclusions**

One of the most successful solution to chain code rotation variability was the introduction of the angular chain code A8. Both, rotational and scale variance was address with solution proposed by Lu [169] and commonly used by other researchers. For rotation invariance, it is based on aligning the major axis of a shape with the  $x$  axis in binary grid descriptors. For scale invariance, the major and minor axes are normalized with respect to a standard axis. In [168],

the authors used the method of Lu and further describe the shapes with fixed perimeter grid boxes to construct chain codes with the same length. The reason was the difficulty with comparing chain codes with different length. It is important to note that: boundary box based on major and minor axes is not necessarily the minimum boundary box; multiple major axes are possible to exist; aligning the major axis with the x-axis without additional considerations does not guaranty complete rotational invariance in 0-360 degrees space.

Based on the literature search and review of available contour and shape descriptors, the chain code (CC) was selected for the application to watermark images. The reasons for that are stated here and justified in the following sections. The summary of criteria for selection of the CC is as follows:

- The CC transform is simple, easy to implement, and fast to compute.
- The CC transform is re-generative.
- The CC transform provides a lossless compression which makes it even more desirable to use.
- The existing (partial) solution of rotation and scale variance of the CC transform can be expanded to provide a complete invariance for the applications to watermark images.
- Translational invariance is possible with a simple recording of the coordinates of the start point of the CCA complete solution to the variability of the CC and provides exact regeneration capability.

Based on the observed characteristics watermark images and watermark imaging processes, rotation, scale, and translation invariant descriptor is mandatory (RST-invariant descriptor). An additional variance also need to be addressed and it is associated with the *recto* and *verso* images of the same watermark. This leads to the development of a RSTV-invariant



descriptor based on F8 chain code which is described next. It is a significant improvement that will expand the use of CC as watermark descriptor and also will allow many new applications.

## CHAPTER 4

### ROTATION, SCALE AND TRANSLATION INVARIANT CHAIN CODE

#### 4.1 Introduction

In digital image processing two main factors have limited the application of chain code descriptors for shapes and object boundaries representation: their scale and rotation variance and the ability to effectively compare two chain codes. In this chapter, a new method is proposed and described that makes possible the representation of a paper watermark with rotation, scale, and translation invariant chain codes (RST-invariant CC).

The property  $P$  is said to be invariant to the transformation  $T$ , *iff* a measurement of the property  $P$  commutes with the transformation  $T$ :

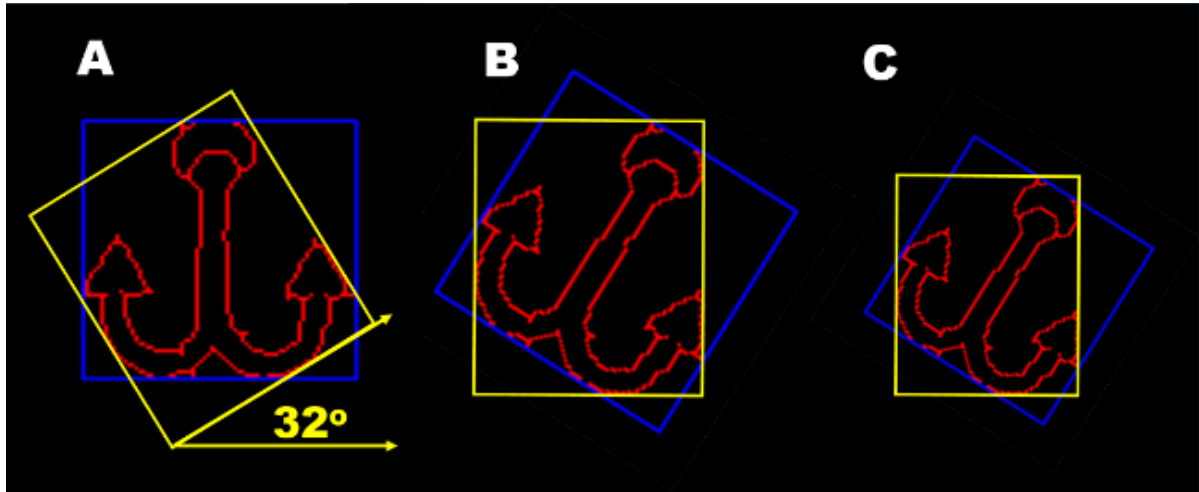
$$P(T(I)) = T(P(I)) \quad (4.1)$$

where  $I$  denotes an image in our case.

In the brief review of available shape descriptors, we discussed the inherent variance with respect to rotation and scale. Our approach to generate RST-invariant CC is based on spatial registration of the watermark image before the CC generation. The rational is always to present to the CC generation algorithm a normalized watermark image with respect to rotation and scale. The normalization is achieved with the application of the circumscribed rectangle around the watermark. Circumscribed means "to draw around" and the official definition of a circumscribed rectangle is the smallest rectangle that contains all of a given set of points in coordinate geometry.

## 4.2 Watermark minimum boundary box (MBB)

In order to normalized an image of a watermark and consecutively generate a RST-invariant CC, we first find the minimum bounding box (MBB) for the binary image of the watermark. An example is presented using the anchor watermark and is displayed in yellow on Figure 4.1A.



**Figure 4.1: Capturing WM relative orientation and scale features (see text).**

First, the circumscribed rectangle of the watermark is aligned with the geometric horizontal and vertical axes (Figure 4.1 in blue color). Next, the minimum boundary box is drawn (Figure 4.1 in yellow color). One may think of the MBB as the rectangle that will produce the smallest possible projection on a rotating axis.

Definitions [163, pp. 626]:

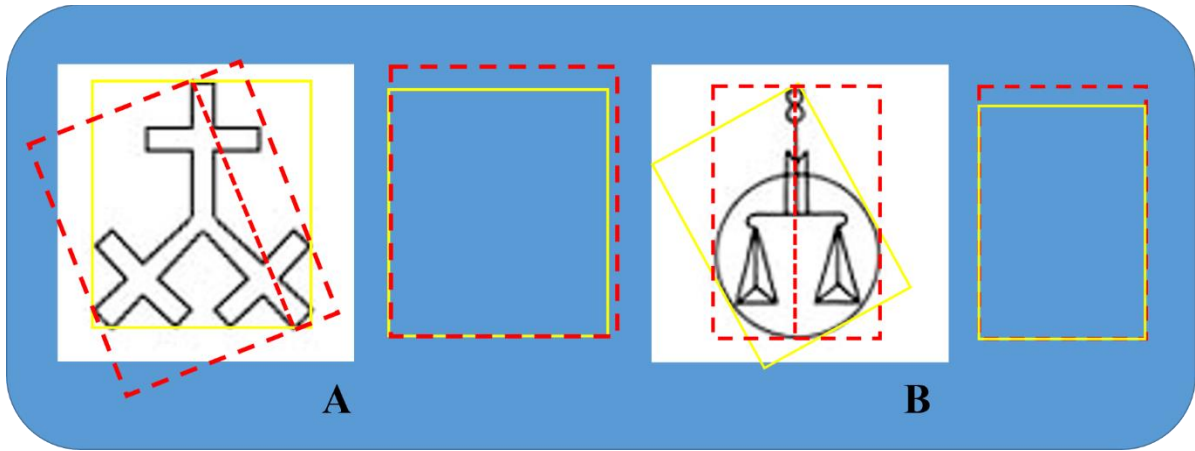
- *Diameter* of a boundary is defined as the Euclidian distance between the two points on the boundary (shape) that are furthest apart.
- *Major axis* of the boundary is the segment connecting the two points on the boundary (shape) forming the diameter.

- *Minor axis* of a boundary is defined as the line perpendicular to the major axis and of such length that a box passing through the outer four points of intersection of the boundary with the two axes completely encloses the boundary (shape).

In general, for irregular boundary (shape), the diameter is not necessarily unique and respectively the major and minor axes.

The definitions are given in order to clarify the difference of the proposed new method for MBB and the rotation and scale solutions currently in use by some researchers in the field of shape and contour description. It is important to point out the difference of the proposed method to the one proposed by Lu [169] and used by other researchers. The new method defines and uses the minimum boundary box while the approach of Lu is to superimpose a rectangle which is aligned with the major axis of the shape. As can be seen in Figure 4.2, the axes defined rectangle is not necessary the minimum boundary box (MBB). In the figure, the major axes of two watermarks are displayed with dashed red lines and the superimposed rectangles that results by following the method of Lu are also displayed with dashed red contours. The yellow rectangles are the MBB for each of the two watermarks. Next to the watermarks, the set of yellow and red rectangles are copied and aligned for visual comparison. Due to the circular element present in watermark B, the superimposed rectangle and the yellow MBB have the same length of their short sides. (Figure 4.2B).

One may think as the MBB in several different ways. It is the circumscribed rectangle (or square in special case) with the smallest area. It is the rectangle (square) with the smallest possible projection on an axis aligned with the short side of the rectangle. On the other hand, the major/minor axes based rectangle is based on maximum and minimum caliper measurements of a shape (boundary).



**Figure 4.2: Comparison of superimposed rectangles. The red rectangles are based on major axis alignment (A). The yellow rectangle is the minimum boundary box drawn by the algorithm of the new method (B).**

It is possible to have more than one major axis and more than one MBB. For the new method, in case of more than one MBB is calculate, the one with the smaller angle of rotation counter-clockwise is selected and used.

### **4.3 Minimum boundary box attributes**

The circumscribed box (SB) is depicted in blue color in Figure 4.1. In the original image the SB has corner coordinates (15, 13) for its top-left and coordinates (87, 87) for the bottom-right corners. This results in a calculated area of 5328 pixels. The yellow color minimum boundary box (MBB) in Figure 4.1, has calculated dimensions of 64 x 77 pixels. Some additional data that could be used is: the area of the MBB (4928 pixels); the short side of the MBB is at 32 degree from the x-axis, measured counter-clockwise, and the long side is respectively at 122 degrees. The CC algorithm can be extended to calculate the area of the shape enclosed by the contour. The anchor's area in this example is 1715 measured in pixels enclosed by the contour. From this data, it is possible to calculate two ratios: (object area / MBB area) and (object area / SB area). In addition, convex hull area can be used to form a

third ratio. This data could be used as differentiating parameters in a richer feature vector, details for which we will introduce later.

#### **4.4 Minimum boundary box for rotation and scale invariance**

Images captured at different relative orientation of the camera with respect to the watermark could be aligned with the support of the MBB. The assumption is the focal plane of the camera is parallel to the watermark plane and the imaging is at the same distance (rotationally variant images of the same size). The binary watermark image can be rotated so that the sides of the MBB are aligned with the  $x$  and  $y$  axes, as shown in Figure 4.1B. Remembering the angle of rotation and then generating the CC, will result in rotationally invariant re-generative representation: the same starting point to generate CC for all watermark images preprocessed with the rotational alignment step.

The scale variability is resolved by scaling the binary watermark image to obtain a size that is the standard for the particular database or catalog, for example short side of MBB 150 pixels (Figure 4.1C). To clarify, in order to match a given watermark by comparing it to the watermarks in a database, before the generation of the chain code, perform the following preprocessing steps:

1. Transform the image to binary;
2. Perform morphological operations including skeletonization;
3. Draw the MBB and obtain its parameters;
4. Rotate the image to position at which the short side of the MBB is horizontal;
5. Scale the image to MBB dimension used for the database.

Of course, some additional provisions have to be implemented in the algorithm. For example:

- If the MBB is a square, make alignment with the smallest possible angle of rotation;
- Always make alignment rotation to position the geometric center of gravity of the watermark closer to the  $x$  axis;
- To resolve recto and versa variability, use horizontal rotation around the axis of the aligned MBB.

#### 4.5 Watermark *recto* and *verso* images

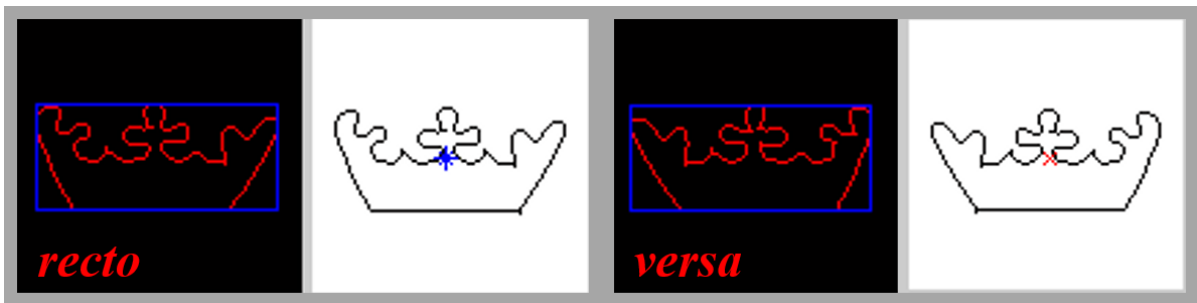
Watermark images have an additional degree of freedom – they can be acquired *recto* and/or *verso*. The rotation and scaling normalization will result in two different images unless the watermark is symmetrical. In general, this is not the case and the processing have to account for the possibility of a “mirror” images (*verso*) of the watermark (*recto*). Similar (but different) relationship exist between the watermark as indented from the wires on the mold side of the paper compared to the watermark viewed from the felt side of the paper sheet.

Figure 4.3 displays four watermarks from the Bernstein collection. They look like two pairs of watermarks with mirror images of each one. A natural question is if they are four different independent watermarks or two watermarks imaged *recto* and *verso*.



**Figure 4.3: Watermarks from Bernstein collection. Are they four separate watermarks or *recto* and *verso* images of two watermarks?**

Figure 4.4 displays the skeleton representation of a Crown watermark *recto* and *verso* images with the minimum boundary boxes superimposed on the images. The two pairs of watermarks are visually mirror symmetrical to each other. One fast way to check for symmetry is to find some additional information. For example, divide the MBB-defined ROI on sub regions and count the non-zero pixels. Even simpler way is to locate the centroids of the images. The images in Figure 4.4 display the binary *recto* and *verso* images (black background) with the MBB. The locations of their respective centroids are marked on the binary skeleton images (white background). The centroid of the *recto* image is marked with blue star. The centroid of the *verso* image is marked with red x sign.



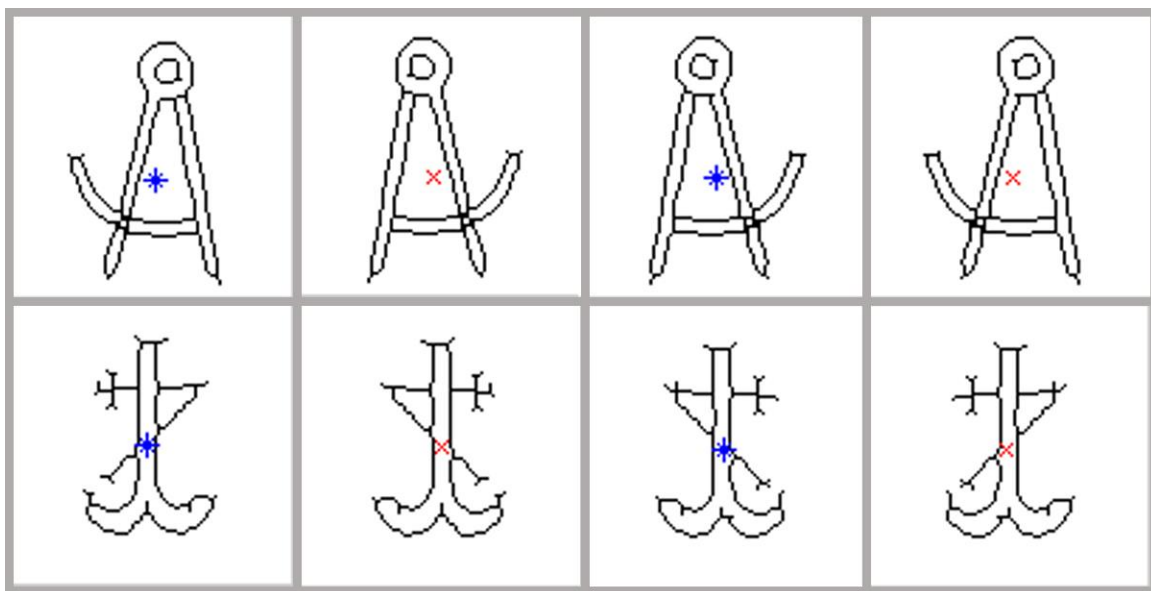
**Figure 4.4: Binary *recto* and *verso* images of Crown watermark with the location of their respective centroids.**

The difference of the centroid location is noticeable. The Y-coordinates of the centroids are the same. The X-coordinates differ by three pixels with respect to the left edge of the MBB in these 100 x 100 pixels images. Exploring other watermarks with similar measurements and observations, the proposed method for resolving the *recto/verso* difference (variability) is to include the centroid or additional measurements into the feature vector description of the watermarks. This approach creates V-invariant CC descriptor. To summarize, the RST-invariance of the CC representation is extended to RSTV-invariant CC representation. In Chapter 5, the subtle visual difference of *recto* and *verso* images of close to symmetry



watermarks becomes very pronounced when their similarity measure is calculated based on the generated chain codes.

The centroid (or other binary image statistics) may be used to test if some of the mirror images in the Bernstein collection are different watermarks or *recto* and *verso* images of the same watermark. Obviously, there are other image processing techniques that can be used. In Figure 4.5, the centroids for the images of Figure 3.4 are displayed overlaid on the skeleton representation of the images. The coordinates of the centroids indicate that the images are with exact mirror symmetry to each other, which is a clue of *recto* and *verso* images of the same watermark.



**Figure 4.5: Centroids of the watermark skeleton images indicating *recto* and *verso* identity.**

The Verso invariance of CC representation is further explore after the introduction of watermarks similarity measure based on CC descriptors.

## 4.6 Conclusions

The inherent variance of shape descriptors including chain code can be overcome by orientation and scale normalization of the shape's image. The approach to generate RST-invariant CC is based on spatial registration of the watermark image before the CC generation. Normalization is achieved with the utilization of the circumscribed MBB of the watermark. This approach has a unique and repetitive solution. Remembering the angle for alignment rotation provides exact reversible restoration.

Resizing the MBB provides scale normalization and a solution to scale variability.

Including additional pixel distribution metrics or the centroid's location, the generated feature vector may account for *recto* and *verso* variability of the watermarks.

Additional provisions in the algorithm are necessary for special case contours and shapes to provide unambiguous CC generation.

More complex contour configurations need iterative description with sequential removal of the described portions of the figure.

The use of MBB results in ability to generate additional features which may be used as differentiating factors for an extended measure of shape (watermark) similarity.

## CHAPTER 5

### COMPARISON OF RST-INVARIANT CHAIN CODES

#### 5.1 Introduction

Ability to compare is based on the ability to measure the degree of similarity (or dissimilarity) with the proper use of an accepted metric or combination of metrics. The more precise and accurate are the metrics, the better is the confidence in the results. For images, the comparison is in a two-dimensional (matrix) space. For images, especially large one, the direct comparison is computationally intensive and respectively slow for larger datasets. The most common approach is to generate a template of the images – a sparse representation of features, and use these templates for comparison of the originals for which they are made.

In general, each image in a dataset is represented by a  $n$ -dimensional feature vector  $F = f_1, f_2, \dots, f_N$ . Then, the similarity between a query image feature vector  $Q = q_1, q_2, \dots, q_N$ , and feature vectors in the database  $F$  is given by a metric measure. Initially, the commonly used similarity measures were tested to evaluate their performance with the CC-based feature vectors. Some of the typical measures which were tested are:

**L1 norm:** The L1 norm is sometimes also called Manhattan or city block distance. The histogram intersection is identical to the L1 norm in case of normalized histograms.

**L2 norm:** A similarity measure between two vectors, based on the Euclidean distance.

**Correlation coefficient:** The correlation coefficient can be interpreted as the dot product of two standard vectors divided by the rank of these vectors.

**Chi-square measure:** The Chi-square similarity measure is based on the well-known chi-square test of equality.

**Matusita distance:** The Matusita distance describes a vector metric that fulfills the Cauchy-Schwartz inequality (generalized triangle inequality).

**Bhattacharyya:** The Bhattacharyya similarity measure [6] is often interpreted as a geometric similarity measure.

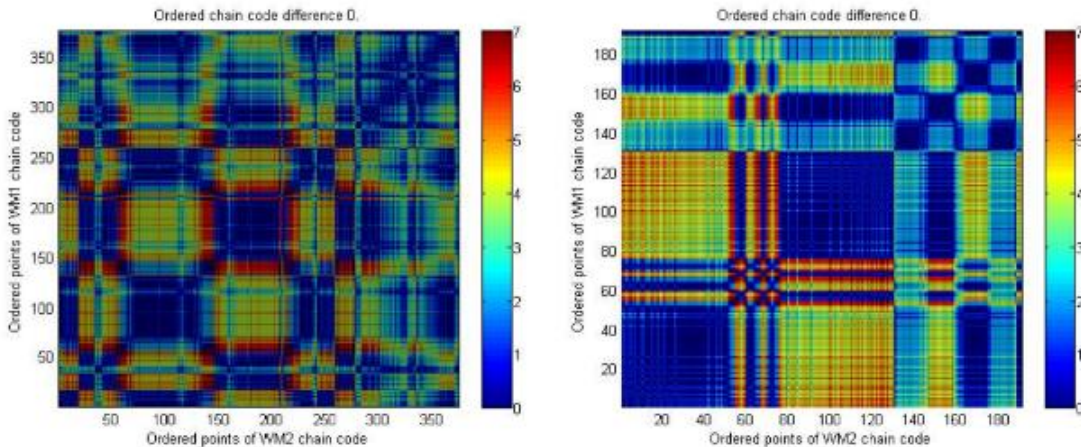
**Probability distributions based metrics** (Kullback-Leibler, Jensen-Shannon divergence, and others) were considered but not explored for testing in this study oriented to comparing the CC feature vectors.

The original testing of existing methods had to resolve the issue of comparing feature vectors of different lengths. As shown, CCs are applicable for shape and contour description. The chain code of a watermark successfully forms a unique feature vector  $F$  in one-dimensional space representation of the two-dimensional watermark image. As described in Chapter 3, one of the main obstacles to successfully use CC as descriptors for identification are their variance to rotation and scale. Another reason is the absence of an efficient method for CC comparison. In the following sections of Chapter 5, additional distance metrics are described. They were reviewed as a candidates for similarity (distance) measures suitable for CC feature vectors. Finally, a newly defined distance metric is described that provides a distance metric for a deterministic comparison of RST-invariant chain codes with any (non-zero) length.

## 5.2 Comparing chain codes with the same length

Next, the capability to compare CCs is demonstrated without detail explanation of the comparison method which is presented in Section 5.4. Figure 5.1 displays a heat-map of the results for comparing watermark CCs to themselves. The two watermarks are *Anker* and *Bell* from the Bernstein collection. The matrix on the left is square with  $375 \times 375$  comparing the

similarity of the CC of the Anker watermark to itself (Figure 5.1, left). The matrix on the right is 191 x 191 comparing the CC of the Bell watermark with a CC length of 191 (Figure 5.1, right). As expected, the distances of self-comparison are with 0 values. Visually the similarity is indicated with the dark blue color of the diagonal pixels indicating no difference.



**Figure 5.1: Distance metrics for CC compared to themselves. The calculated differences have 0 values. The similarity score is the sum of the values on the diagonal.**

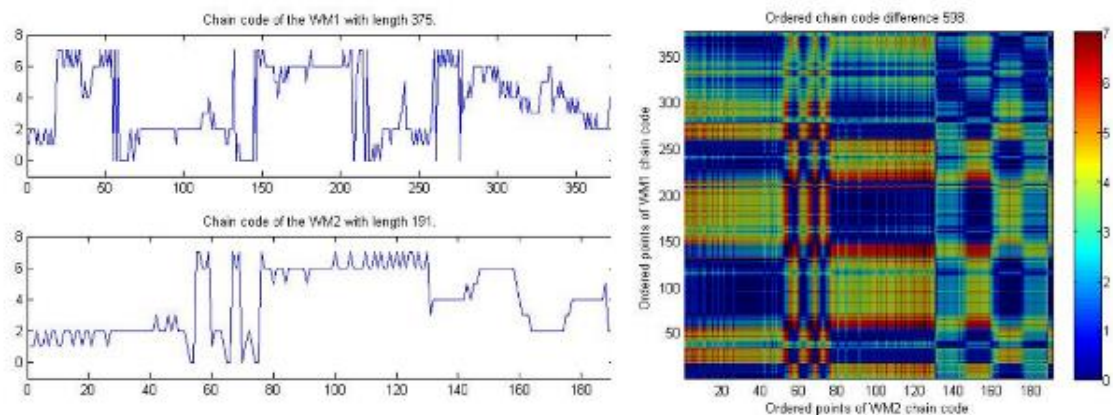
Figure 5.1 demonstrates two examples of the ability to successfully compare watermark CC to itself. In self-comparison the two CC are inherently the same length and respectively the heat-map matrices are square and symmetric. This is a very special and not very interesting case. However, it is important that the result of self-distance is 0.

In the next section, the ability to compare CCs of different watermarks and with different length is demonstrated with an example followed by a description of the development of the CC distance metric in Section 5.4.

### 5.3 Comparing chain codes with different length

When the 4-directional CC is used, one may apply distance metrics develop for the comparison of the amino-acid sequences of DNA and RNA. The use of 4-directional CC is

limited to binary images guaranteed to have neighboring pixels in horizontal and vertical direction, which is not the general case. For 8-directional CC, the used methods are based on string comparison with different cost function for replacements to match. Histogram and different statistics of the CC are also used. After review of applicable algorithms, we develop a distance metric similar to template matching comparison in biometrics. For example, when the templates of test images are compared to the template in the database, the formed square matrix would have minimal values on its main diagonal. Needless to say, comparing a set of templates to themselves would produce a matrix with all zeros on the main diagonal. For the CC, we use similar approach. The example illustrated on Figure 5.2 is for comparison of the CCs of two watermarks: *Anker* with length 375 (top graph) and *Bell* with length 191 (bottom). Visually, the two chain codes look very different. The distance metric computation is presented with a heat map matrix with 375 rows and 191 columns. The sum of the values on the diagonal is 598 and indicates the degree of difference (dissimilarity) of the two CCs and respectively the two watermarks they represent.



**Figure 5.2: New Distance metric for CC comparison. The similarity score of 598 is the sum of the values on the diagonal.**

#### 5.4 Chain code distance (CCD) as a measure of watermark similarity

Chain code (CC) is easy to generate and its regenerative property provides a true restoration of the original contour. When the coordinates of the start pixel are recorded, the restored image is exact. These attributes combined with the proposed solution for the RST-invariance, make the CC an attractive candidate for a contour descriptor. In addition, the *recto* and *verso* variability of watermark images was also addressed. For a complete success, one more problem needed to be resolved: development of a method for comparison of chain codes.

In general, the CCs are of different length. Currently, algorithms for quantitative comparison of chain codes of different size do not exist. As previously noted, many authors have developed and used approaches to resolve the problem with rotation and scale variance of the chain codes and with the difference in length. Histograms of chain codes have been applied with limited success.

Different measures of distance or similarity are applicable for different types of analyses. Multiple distance measures have been implemented in different programming languages and are readily available. Researchers continue to improve the efficiency of complex algorithms which make them applicable to large data sets. The characteristics of the data define the selection of the measuring distance. General and mixed data similarity is typically measured by feature-derived distances. Chain codes can be viewed as features generated from the watermark images. The watermark can be compared with image distances or with geo-spatial distance measures. For images, information theoretic and probabilistic based measures are used [226-227]. Spatial distance based measures generally consider the pairwise distances between the compared point sets. Examples are the average distance (the average of all pairwise distances) and the Mahalanobis distance, which compares estimates of the point sets as two

hyper-ellipses. The methods are not sensitive to the positions of the individual points. A measurement that takes into consideration the spatial position of each individual point, is the Hausdorff distance - a max-min distance, which makes it capable of considering the spatial properties in the measurement, e.g., the boundary of an object (contour). However, in some 2-dimensional applications, this makes it sensitive to outliers [227, 228].

When viewed as a sequences of numbers, there are many distance measures for numeric data that can be applied. For numeric data in general, the following distance (similarity) measures are commonly used: Manhattan, Euclidean, Squared Euclidean, and Normalized Squared Euclidean, Canberra, Cosine, Correlation, Binary, Bray-Curtis, Chessboard, Warping, Canonical warping, and others. Viewed as strings, the similarity of chain codes can be expressed with available string data comparison algorithms: Damerau-Levenshtein distance, Smith-Waterman similarity, Hamming, Needleman-Wunsch similarity and some others. In this regard application of sequence alignment and sub-sequence positions have to be used as well.

Additional distance measurements were considered for development CCs comparison: Hausdorff distance, Frechet distance, Jaccard distance (similarity of asymmetric binary attributes), difference with simple matching coefficient (market basket analysis), Tanimoto distance (a method of classification based on a similarity ratio, and a derived distance function), Marczewski-Steinhaus distance (more generalized Jaccard distance), and others [229-231]. The search was to find the most applicable distance measure suitable to the characteristics of the CCs. More specifically, CC are in general with different length and the length itself cannot be normalized. In pair-wise comparison, the length of the shorter chain could be extended to match the length of the longer chain. CC has a limited space presentation (8-character alphabet) which is known or even can be selected. For example, negative and



positive decimal values have been considered. Using alphabet  $\{-4, -3, -2, -1, 1, 2, 3, 4\}$  and absolute values for distance measure was explored and it is still an option.

Many of the listed algorithms were considered and some of them were tested to measure the difference (similarity) of the chain codes generated from the binary images of the watermarks. While the difference in length could be viewed as a complication, several characteristics of the chain codes facilitates their comparison: they are a continuous sequence of positive real integer numbers and their numeric values are in *a priori* known space from 0 to 7 only: bounded (normalized) data with no outliers.

The directed Hausdorff distance  $H$  between two point sets  $A$  and  $B$  is the maximum of distances between each point  $x \in A$  to its nearest neighbor  $y \in B$ . That is:

$$H(X, Y) = \max_{x \in A} \left\{ \min_{y \in B} \|x, y\| \right\} \quad (5.1)$$

where  $\|\dots\|$  is any norm e.g., the euclidean distance function. The characteristics and capabilities of Hausdorff distance measure makes it a good candidate for similarity measurement of CC. However, an examination of its definition reveals that  $H(X, Y) \neq H(Y, X)$  and indicates that the directed Hausdorff distance is not symmetric. For example,

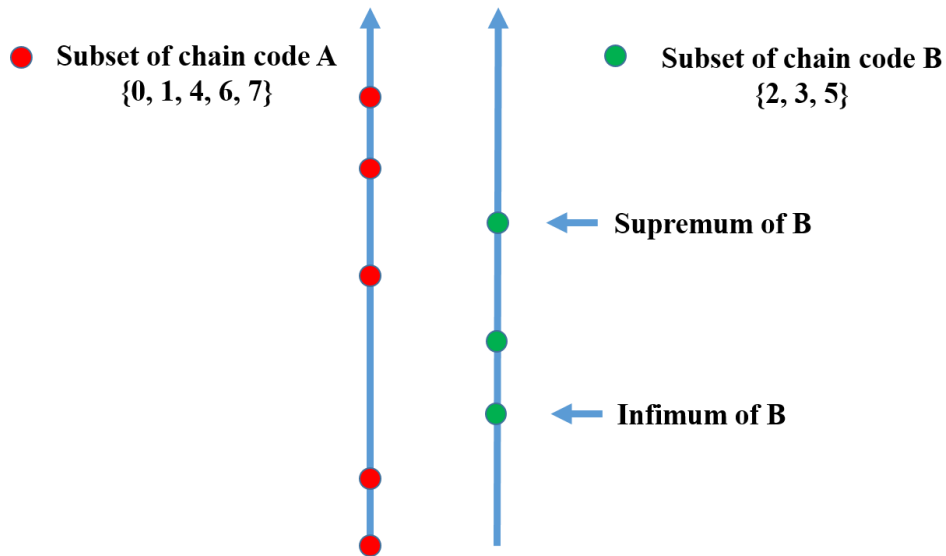
$$H([1,3,6,7], [3,6]) = 2 \text{ and } H([3,6], [1,3,6,7]) = 0 \quad (5.2)$$

For the intended matching comparison of one-to-many chain codes of watermarks, the comparison is in one direction: one CC to the CCs in a database. Nevertheless, it could be a major problem as illustrated by the example (5.2). The problem is solved by using the maximum value of the directed Hausdorff distances pair-wise in both directions, is an easy way to make the Hausdorff distance symmetric:

$$H(X, Y) = \max\{H(X, Y), H(Y, X)\} \quad (5.3)$$

and remove the constrain of directional comparison difference.

More formal definition of the Hausdorff distance is given using *infimum* and *supremum* of two datasets comparison [233, 234]. Figure 5.3 is a simple illustration for two short chain codes of five and three data points.



**Figure 5.3: Supremum and infimum of two sets A and B.**

Hausdorff distance was developed and introduced by the German mathematician Felix Hausdorff [232]. It is also called Pompeiu–Hausdorff distance and is used to measure how far (in a metric space) two subsets are from each other. In image processing, one application of the Hausdorff distance is to find a given template in an arbitrary target image. The template and image are often pre-processed with an edge to make a binary representation. Every point in the binary image of the template is treated as a point in a set, the template "shape" set. The binary target image is treated as a set of points. The algorithm then tries to minimize the Hausdorff distance between the template and some area of the target image. The area in the target image with the minimal Hausdorff distance to the template, can be considered the best candidate for locating the template in the target [233].

By definition, the Hausdorff distance can be considered the longest distance one can be forced to travel from a point in one of the two sets, to the other set. In other words, it is the greatest of all the distances from a point in one set to the closest point in the other set. That is a great characteristic for a distance measure: it gives the worst case similarity measure.

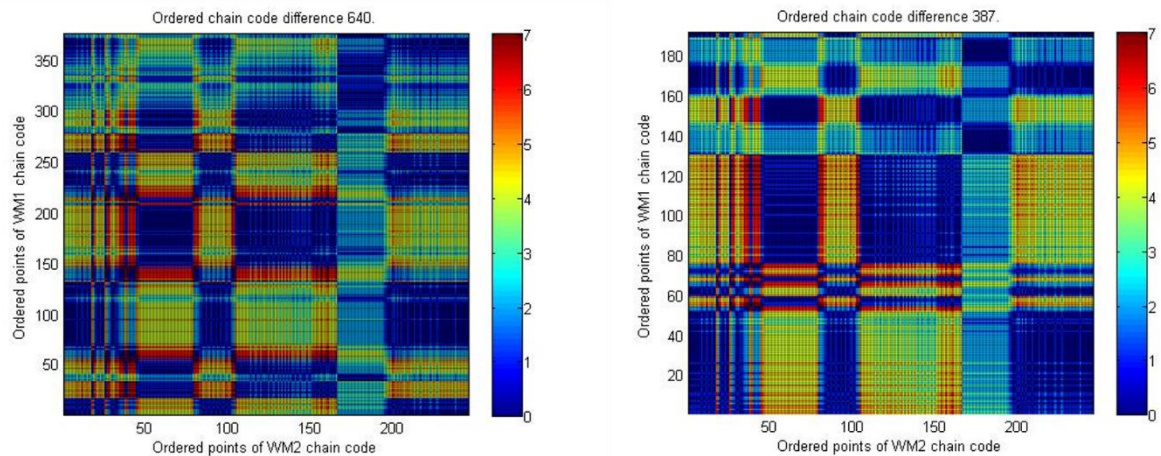
The Hausdorff distance is used in many applications and have many implementations. Researchers have noted the importance to optimize the runtime and memory required for large datasets and how the performance vary in relation to the sets size, grid size, density, sparsity, and generality of the data [234, 235]. Fortunately, chain codes are a one-dimensional datasets which are also relatively small size (short length), equal density and without missing values (no sparsity). Unfortunately, Hausdorff distance cannot be used directly for the comparison of two chain codes. The CCs are bounded in  $[0, 7]$  space and two CCs are always overlapping in this space resulting in 0 Hausdorff distances. In next section, a modification of Hausdorff-like metric is presented which can be successfully used for comparison of datasets in very small bounded space.

## **5.5 Application of CCD for watermarks comparison**

To evaluate the similarity of two CC, first we form replicas and tile the elements of the two CC. This allows the comparison of two arrays with different lengths. For lengths  $L_1$  and  $L_2$  this forms  $L_1 \times L_2$  pairs of distances. The distance metric calculation is performed for every pair of values as the square root of the sum of the squared value differences at each corresponding index. The maximal minimum distance per index is retained traversing forward. As mentioned previously, to assure symmetry of the distance measure, the calculation is

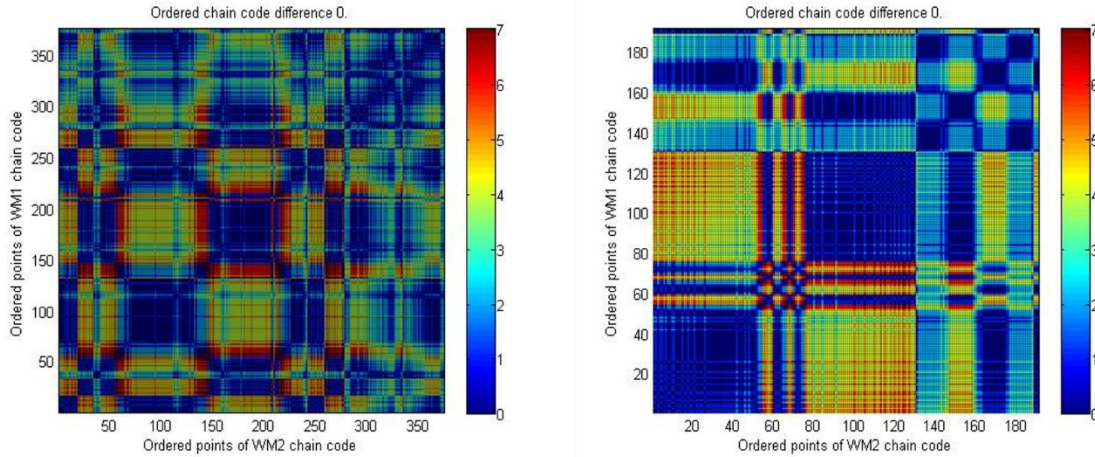
performed traversing the array in reverse. The overall distance is the square root of the maximum of the two (forward and reverse) passes of the arrays.

Figure 5.4 provides two more examples for comparison of watermarks based on their chain codes and chain codes distances. The heat map on the left is showing the CC differences (0 to 7) at every index for watermark *Anker* and watermark *Hand*. The Chain code distance measure (similarity score) is 640 indicating high difference. The heat map on the right demonstrates the index differences between watermark *Bell* and watermark *Hand*. The overall distance difference score is 387.



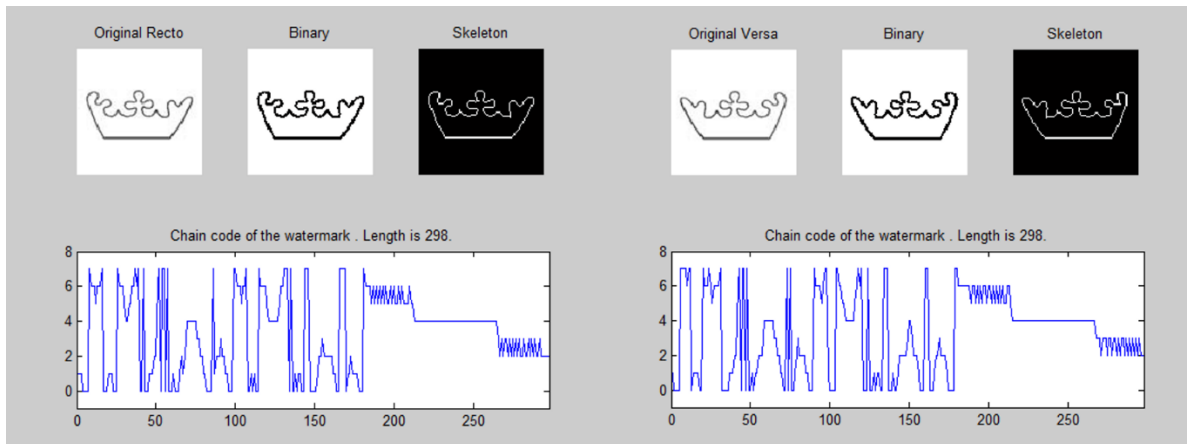
**Figure 5.4: Comparison of watermarks based on their chain codes distances. The CC difference for watermarks *Anker* and *Hand* is 640 (left). The difference score between watermarks *Bell* and *Hand* is 387 (right).**

Figure 5.5 shows the results for self-comparison (a test for match) for two watermarks. On the left, the chain code of the watermark *Anker* is compared to itself. On the right, the chain code of the *Bell* watermark is compared to itself. As expected, the chain code distance measure is zero indicating a perfect match.

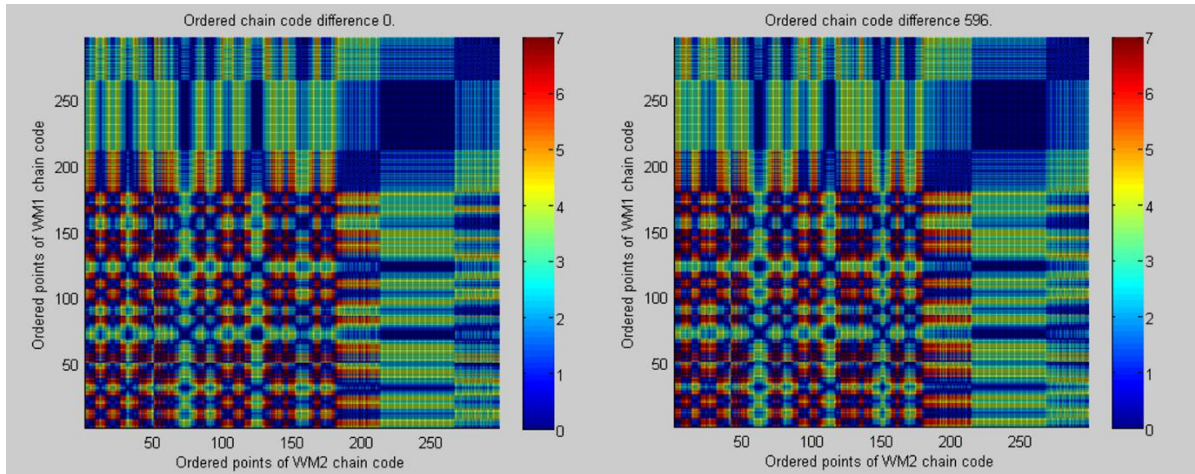


**Figure 5.5: Chain code self-comparison for the watermark *Anker* (left) and for the watermark *Bell* (right). The difference measure is zero indicating a perfect match.**

Figure 5.6 displays the recto and versa images of Crown watermark and the CCs generated from the images. Visually, the images and the chain code patterns look similar. The perceived similarity is rejected by the application of the CCD metric. In Figure 5.7, the heat map on the left displays the self-comparison of the CC representing the recto image. On the right side is the heat map display for CC comparison of *recto* and *verso* images. The perceptually similar CC in Figure 5.6 are actually very different. The calculated distance value is 596.

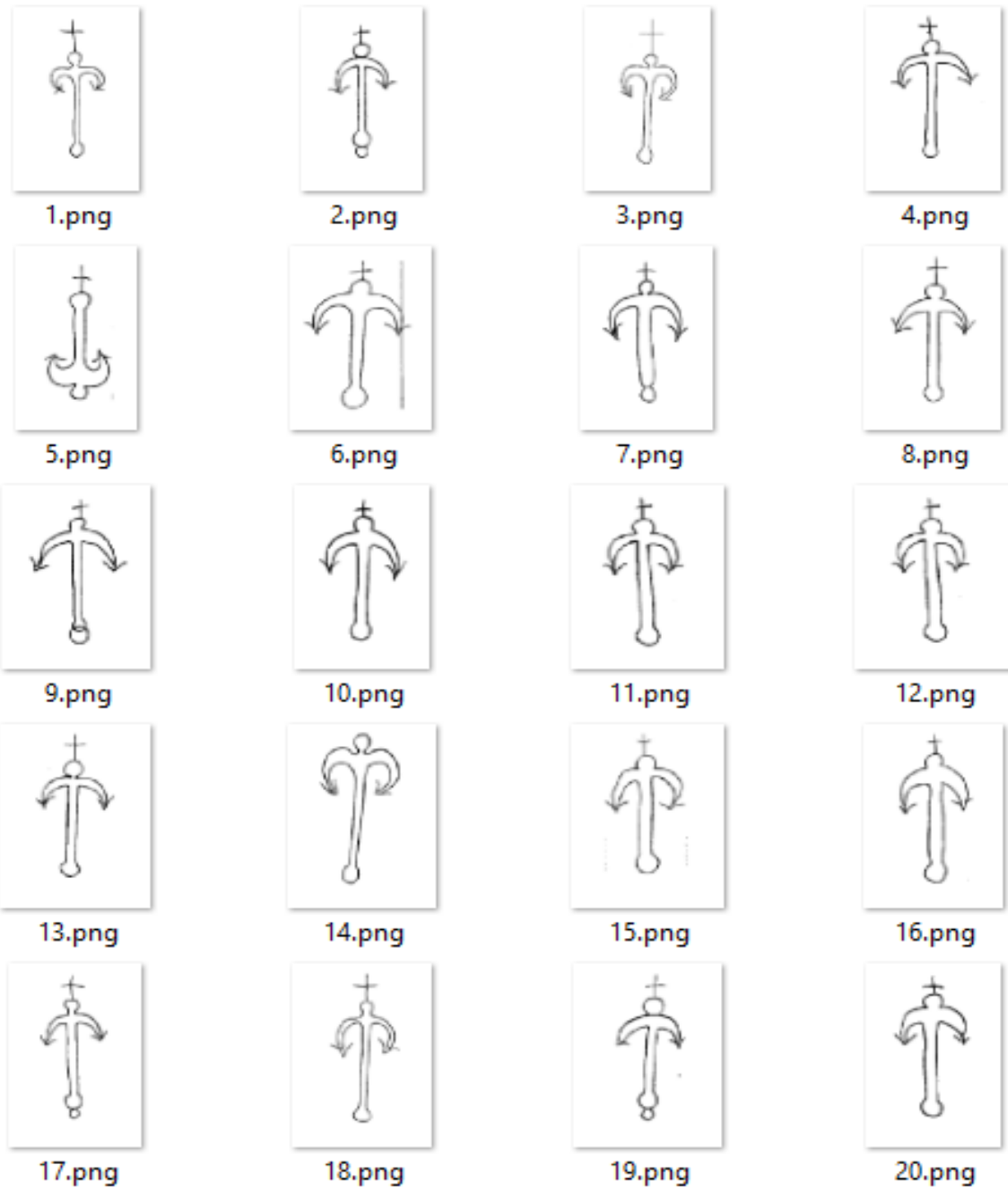


**Figure 5.6: Chain code of Crown watermark *recto* (left) and *verso* (right). The Chain code have the same length and similar pattern.**

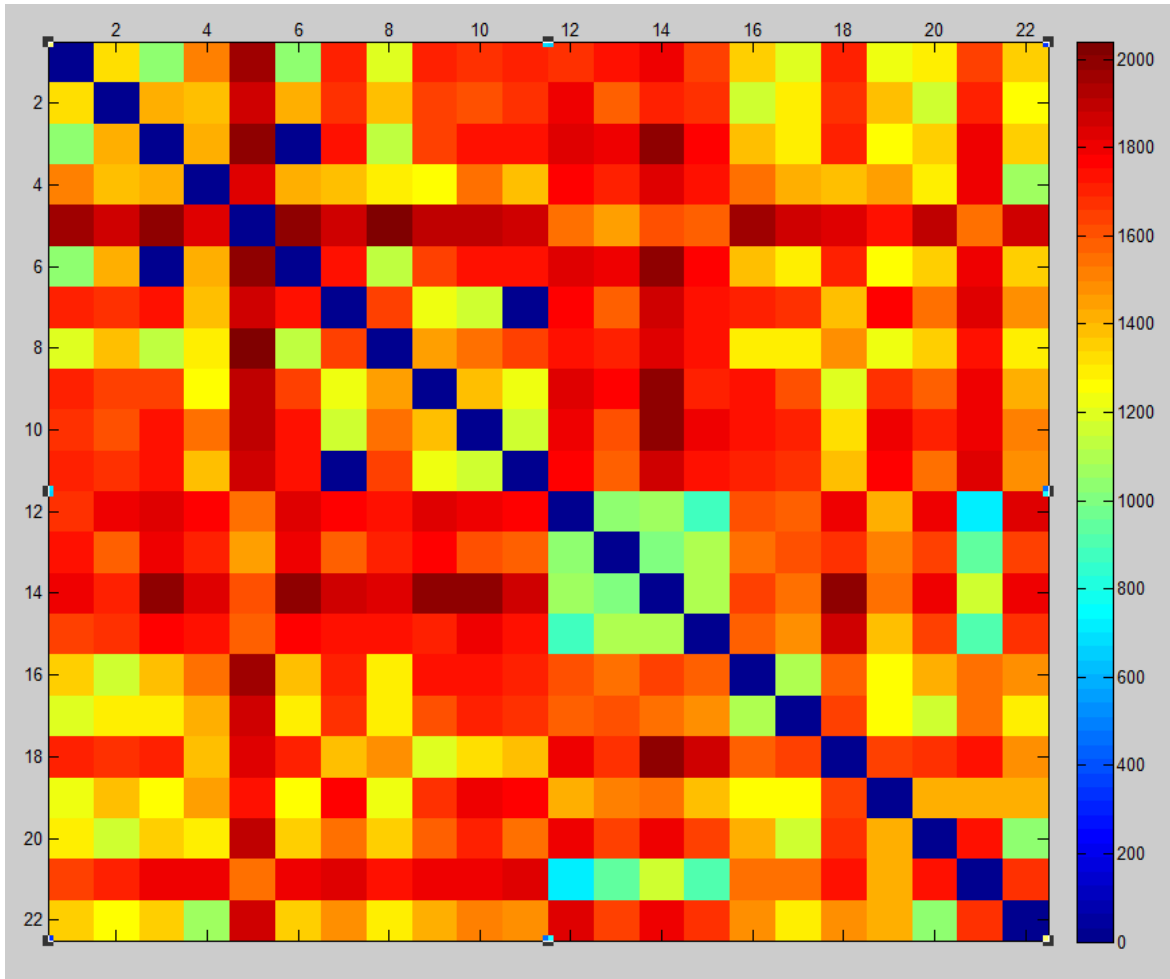


**Figure 5.7: Crown watermark *recto* and *verso* comparison. Self-comparison CCD is 0 (left). The *recto* vs. *verso* CCD is 596.**

Next, the resolving capability of the new method is demonstrated with a set of watermarks from the same class. Figure 5.8 shows a set of 20 watermark images from the class *Anchor*. Visually, the watermarks appear similar, nevertheless different. It is interesting to observe that most of the anchors show the same orientation except the first image in the second row. The image reveals that the anchors are featured with a tiny cross atop and with cusp-like structures at the outer endings. The *Anchor* was a very popular watermark and for differentiation additional motives were added with time. The class *Anchor* possesses a large intra-class variation of shapes, i.e. many anchors have no crosses or show very different endings. Those characteristics are described in many of the datasets with anchor watermarks. The resolving capability of the chain code distance measure is tested using the class *Anchor* watermark images. Two images are selected at random, copied and added to the set of 20 images to form two duplicate pairs. The duplicates are used to test the matching (discriminating) fidelity of the CCD. Figure 5.9 displays the distances between the chain codes for the original 20 images plus the two duplicates.



**Figure 5.8: Watermark images of class Anchor.**



**Figure 5.9: Chain codes distances for class Anchor watermarks. Two watermark images are copied and added to the original dataset for testing. They have formed two matching pairs (3 & 6) and (7 & 11).**

Figure 5.9 displays the CC based distances for 22 watermark images (20 original watermark with two duplicates added). After binarization, skeletonization, RST-invariant transformations, and CC generation, the total number of 22 CCs are compared. The self-comparison results in 0 distances (blue color diagonal). The first duplicate forms a matched pair with a 0-distance at indexes 3 and 6 and (7&11). It is obvious that if the images are sorted in ascending order of their CCD, it will result in the complete similarity ranking for the images.

The newly developed CCD also provides the ability for discrimination or identification of *recto* and *verso* watermark images. Figure 5.10 displays four watermarks from the Bernstein

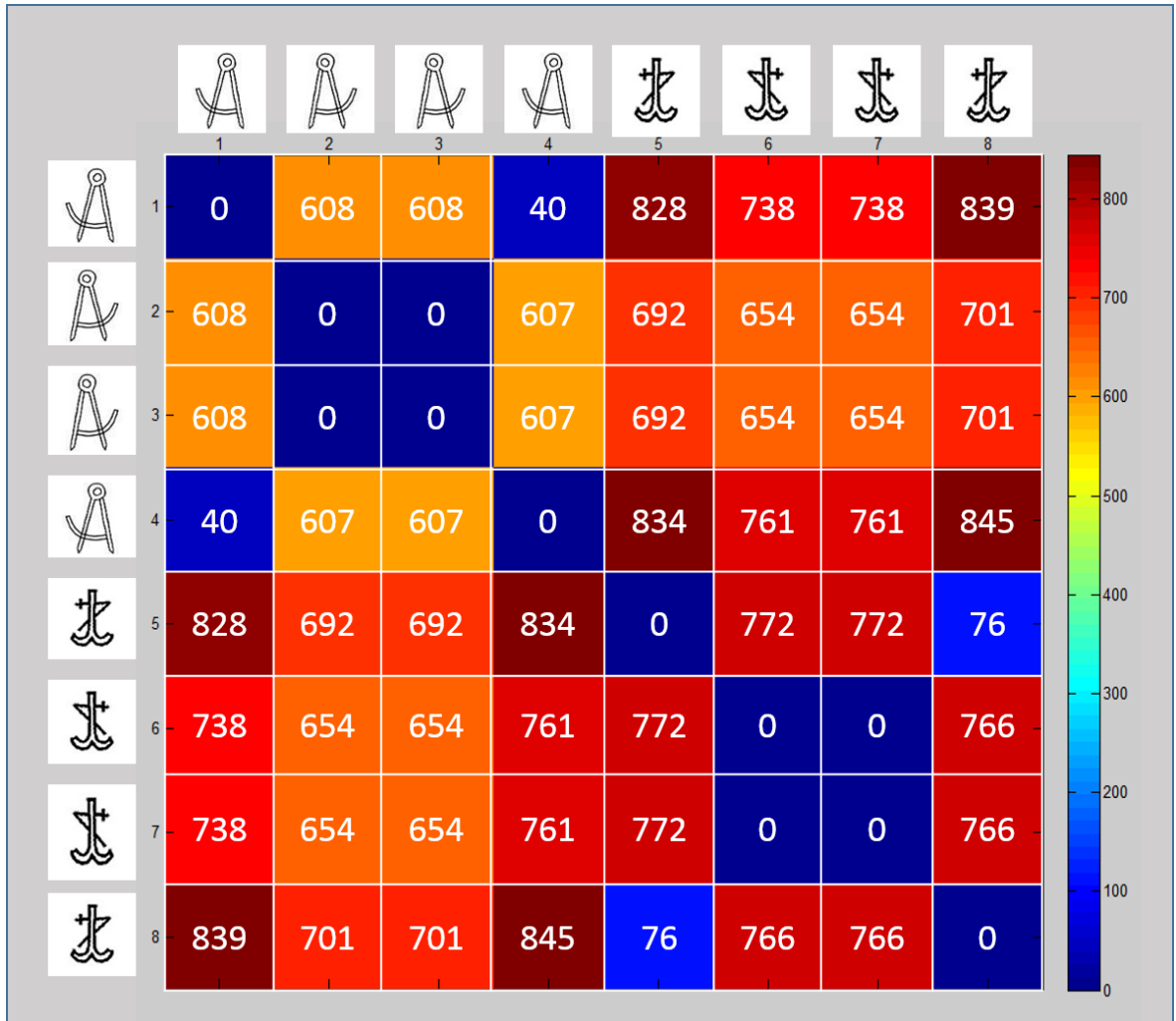


collection with their horizontally flipped images. Those are the same watermark images which were previously displayed in Figure 4.3.



**Figure 5.10: The watermarks from Bernstein collection from Figure 4.3 with their mirror images.**

The application of the CCD generates the 8 x 8 similarity (difference) matrix which is displayed as heat map in Figure 5.11 with the calculated distance measures.



**Figure 5.11: CCD for the four watermarks and their mirror images.**

In the matrix, image at position 2 (row and column) is the mirror reflection of image at position 1 (row and column). The image at position 4 is the mirror reflection of image at position 3 and so on for the next images. The calculated chain code distances are printed on the heat map representation of the symmetric CCD matrix. The CCD values generated for the mirror image of watermark 006003026 and watermark 006003027 at indexes (2, 3) and (3, 2) are 0. The CCD values generated for watermark 006003026 and the mirror image of watermark 006003027 at indexes (1, 4) and (4, 1) are 40, indicating very close similarity. The distance differences of 0 and 40 is leading to the following (hypothetical) conclusion: Watermarks 006003026 and 006003027 of the Bernstein collection are the same watermark represented with *recto* and *verso* images. Or they are well made twins, with the second figure attached flipped to the chain and laid wires. This is less likely but not impossible.

Similar conclusion can be made for watermarks 010001002 and 010001003 of the Bernstein collection. The distance value is 76. Close examination of the skeleton images of the original and the mirror images prompts to investigate further the dependence of the CC on the skeletonization and, in general, the image preprocessing.

## **5.6 Conclusions**

In Chapter 5, the development of new method for generation of RST-invariant chain codes was presented. The method provides capability for normalization of watermark images and therefor, accurate comparison of their generated chain codes (CC). The results prove the efficiency of CC-based templates for watermarks comparison. The RST-invariant CCs are lossless compressed representation of the watermark images with very high ratio of compression. Many possible distance metrics were reviewed as candidates for watermarks chain codes comparison. The implementation of a Hausdorff-like distance metric, modified

with respect to the chain codes characteristics, is described. It was demonstrated that the new distance metric performs efficiently and accurately.

The CCD can be used for measuring the distances of the chain codes and respectively the similarity of watermarks.

## CHAPTER 6

### AUTOMATIC IDENTIFICATION OF WATERMARKS WITH CHAIN CODES

#### **6.1 Introduction**

Identification is based on the ability to successfully compare an object of interest to a set of other objects. In identification systems, the user makes an implicit “one-to-many” search of the entire enrolled database. When the degree of similarity is sufficient, the object is identified. When the comparison does not result in a satisfactory outcome, the object is considered not present in the space of comparison. For images, a direct comparison is in a two-dimensional (matrix) space. For CC we compare in a one-dimensional (vector) space. As described in Chapter 3, the compression that CCs provide simplifies the comparison for matching process. However, one of the main obstacles to successfully use CC as descriptors for identification are their variance to rotation and scale. Another reason is the absence of an efficient method for CC comparison. In the following sections of Chapter 6, an approach is described that provides the performance testing results for the deterministic comparison of chain codes with any length. The test is performed with 250 watermarks dataset. For comparison, additional tests were conducted using modern image classification methods with deep learning convolutional neural networks and bag-of-visual-words.

#### **6.2 Watermark datasets**

An important note regarding the use of terminology is the referral to a collection of images as a “corpus”. Any additional information about the collection of images and how they are produced is referred to as the “database”. In this thesis, the preferred and most frequently used

term is “dataset” as related to the set of images only or the corpus. Using the corpus, CC representation of the watermarks are generated. The constructed searchable space is sufficient for one-to-many comparison and matching of the generated CCs. However, a relational connection with the corpus is necessary to maintain CC-to-watermark relationship.

### **6.3 Performance evaluation**

The initial steps in an evaluation study are to decide what exactly the performance results are trying to determine. Scholars agree that there are at least three evaluation types: technology, scenario, or operational evaluation [236]. For the case of watermarks identification the goal is to demonstrate two aspects:

1. Chain codes correctly, uniquely and persistently describe the watermarks.
2. The newly developed distance metric based on the comparison of CC (similarity measure) accurately matches (identifies) a watermark if it has been enrolled in the relational database.

“Accurately identifies” is interpreted as a performance with an acceptable degree of errors for a given number of comparisons. In the case of watermarks search and retrieval, the cost of false positive or false negative is the degree of annoyance to the user of the system.

For watermarks available from datasets (databases), choice of the performance evaluation type, determines the technology evaluation approach with respect to the CC-based matching algorithm. The technology involved in the acquisition of the watermark images is not evaluated and a simplified assumption is that the images in the database are of similar and sufficient quality. Because the generation of CC is preceded with morphological preprocessing of the watermark images (binarization and skeletonization), the performance metrics indirectly

evaluate the fidelity of CC generation as well. When a system for identification of watermarks based on CC is installed, it has to be tested with **operational** and/or **scenario type** evaluation.

In this thesis, the focus is on **technology type** evaluation of the proposed new method for automatic identification of watermarks based on CC.

### 6.3.1 Dataset sample size

In any study that involves evaluation of an algorithm, the size of the sample size is an important factor for the statistical significance of the results. Design of a good experimental work always includes the process of sample size selection. The sample size defines how accurately the performance errors can be measured. The larger the test sample size, the more accurate the results will be and with a better confidence level.

The identification of watermarks is an implicit matching process. The match/non-match result is based on many pair-wise comparisons, i.e. one-to-many search for match. Before proceeding with the justification of the sample size for statistical significance, several definitions are presented.

Confidence interval: Very often a 95% confident interval is used. For a parameter  $x$  that estimates a lower estimate  $L$ , and an upper estimate  $U$ , such that the probability of the true value being within the interval estimate is the stated value (e.g.  $P(x \in [L, U]) = 0.95$ ). The smaller the test size the wider the confidence interval  $[L, U]$  will be [237-239].

When an algorithm with an observed performance close to 100% accuracy (respectively, close to zero probability of errors), it is challenging or impossible to use the available statistical test for sample size calculation. The formulae with  $\mathbf{p} * (\mathbf{1} - \mathbf{p}) = \mathbf{0}$  term cannot be used. For these cases, the statisticians use the “Rule of 3” [238]. The rule is used to estimate the lowest

error rate that can be statistically established with a given number of (independent identically distributed) comparisons. This rule  $p \approx 3/n$  defines the value of the error rate  $p$  for which the probability of zero errors in  $n$  trials, is purely by chance. For a sample size  $n$ , with probability  $p_0 = 0$  the interval from  $0$  to  $3/n$  is a 95% confidence interval for the rate of occurrences in the population.

For a pair-wise matching with a success rate of 100%, the measured accuracy means that a negative result has not been observed to occur in  $n$  Bernoulli trials. In the case of CC comparisons, the matching process has only two possible outcomes (match/non-match) and can be considered a Bernoulli trial (or binomial trial). For example, if the sample size is  $N = 100$  the  $n = N! / (2 * (N - 2)!) = N * (N - 1) / 2 = 4950$ . Usually, a 95% confidence interval is used for the probability  $p$  of an event occurring for a randomly selected sample set in a given population. If we note the number of events by  $X$ , we therefore wish to find the values of the parameter  $p$  of a binomial distribution that give  $\Pr(X = 0) \geq 0.05$ .

The rule can then be derived either from the Poisson approximation to the binomial distribution, or from the formula  $n * (1 - p) = (1 - p)^n$  and by taking logarithms and keeping only the first term of the series expansion of the natural logarithm [240]. In short, using Taylor's theorem the factor of three is for  $n * \ln(1 - p)$  we apply  $\ln(1 - p) = -\ln(0.05) = \ln(20) = 2.9957 \approx 3$ . So  $n * p = 3$  or  $p = 3/n$ . For a sample size of 100, the result is  $p = 3/4950 \approx 0.0006$ . Or for easier reference, given a measured performance with zero errors ( $p_0 = 0$ ), the statistically claimed performance with 95% confidence is that 1 in 1650 comparison trails may result in non-correct match. From  $p_0 + p_1 = 1$ , one could expect for only successes ( $p_1 = 1$ ) the confidence interval to be  $[1 - 3/n, 1]$ .

**Conclusion:** Based on this reasoning using the rule of three with 95% confidence and [0.9994, 1] confidence interval the measured accuracy could be extended to an error of 1 for 1650 comparisons.

**The reverse problem:** Using the rule of three, calculate the sample size needed for a study with 95% confidence and a desired calculated accuracy of 0.999 (0.1% error) given a measured (claimed) accuracy of 1 (100%). What is the required minimum sample size?

**Solution:** From  $0.001 * n = 3$  we get  $n = 3000$ . If we have a sample size of 78 we perform  $78 * 77 / 2 = 3003$  comparisons. To test and demonstrate the claimed accuracy we would need 78 samples and could extend the performance accuracy to 0.999 with a 95% confidence.

To summarize this topic, for confidence level of 90% it is the “rule of 2.3” and for 99% confidence level, we need to use the “rule of 4.6”. Using the “rule of 4.6” with 99% confidence level and 0.1% error, the result is  $n = 4.6/0.001 = 4600$  trials. For that many trials the minimum sample size is 97 samples ( $97 * 96 / 2 = 4656$ ).

The “rule of 3” is used to estimate the lowest error rate that can be statistically established with a given number of (independent identically distributed) comparisons for the most commonly used confidence level of 95%. The identical distribution and independence of the CC generated for the watermarks in the dataset are open for discussion. Fortunately, in statistics,  $n$  greater than 30 is considered a threshold above which the rule of three is a good approximation to results from more sensitive tests.

### **6.3.2 Receiver operating characteristics**

Receiver operating characteristic (ROC) curves are an accepted method for summarizing and visually presenting the performance of a detection system (radar, diagnostic test, pattern

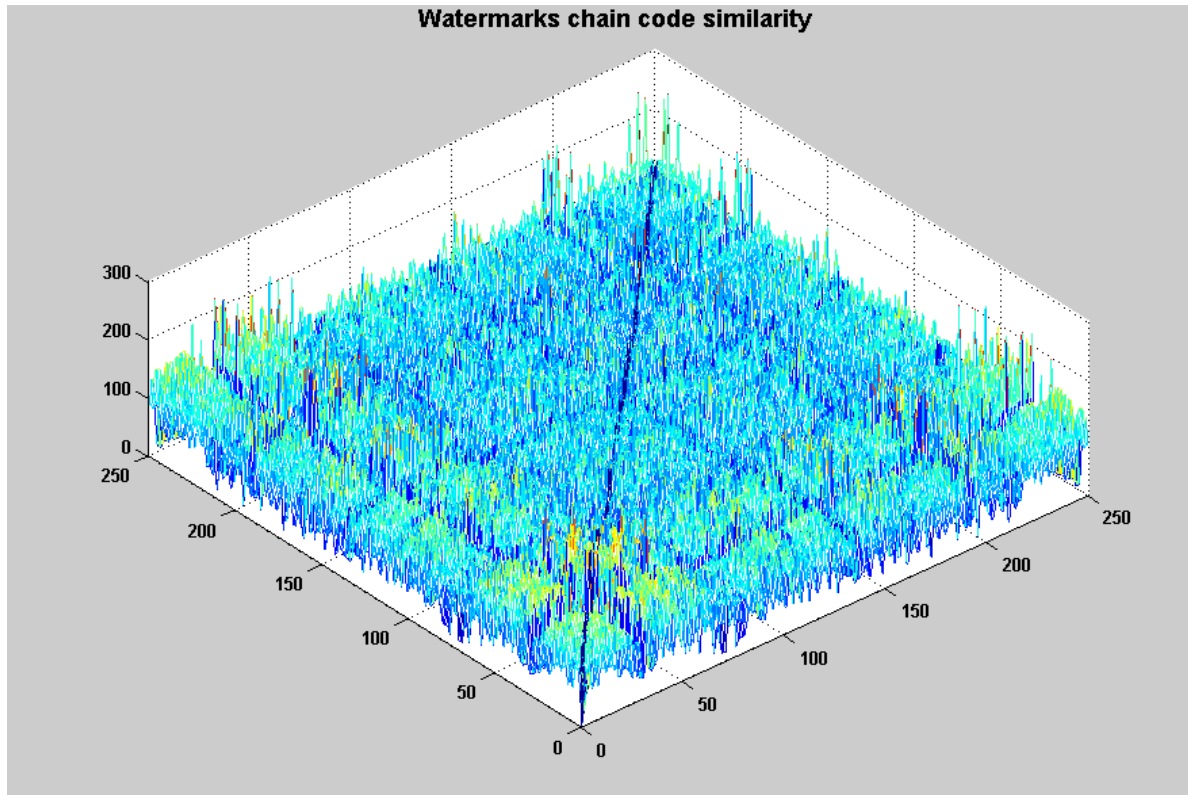


matching and others). An ROC curve plots, parametrically as a function of the decision threshold, the rate of “false positives” (i.e. impostor attempts accepted) on the x-axis, against the corresponding rate of “true positives” (i.e. genuine attempts accepted) on the y-axis. ROC curves are threshold independent, allowing performance comparison of different systems under similar conditions, or of a single system under differing conditions. Figure 6.2 in the next section shows an example of ROC plots. The area under the ROC curve (AUC) indicates with a single value the overall performance of the detection system. A value of 1 represents a 100% accuracy. A value below 0.5 is below the theoretical level of random processes. Obviously, AUC with values close to 1 (or 100%) are desirable because they indicate a well-designed and well performing system (algorithm).

### **6.3.3 Experimental results using CCD**

Next two figures present a summary of the matching performance of the generated CC. Figure 6.1 displays the normalized distances in 1 to N comparison for 250 WM from the Bernstein catalog collection [18]. Each image is preprocessed as described previously (binarization and skeletonization) and the generated CCs are stored in binary files with the same name as the original watermark images. The dataset images are presented in Appendix A.

The ROC curve analysis for the CCs comparison presented in Figure 6.2, indicates an area under the curve (AUC) value of 0.99971 and the equal error rate (EER) is 0.000289. The results are from the comparison of the CC of watermarks images of the dataset using the described CCD metric. The ROC is presented in blue color on Figure 6.2.

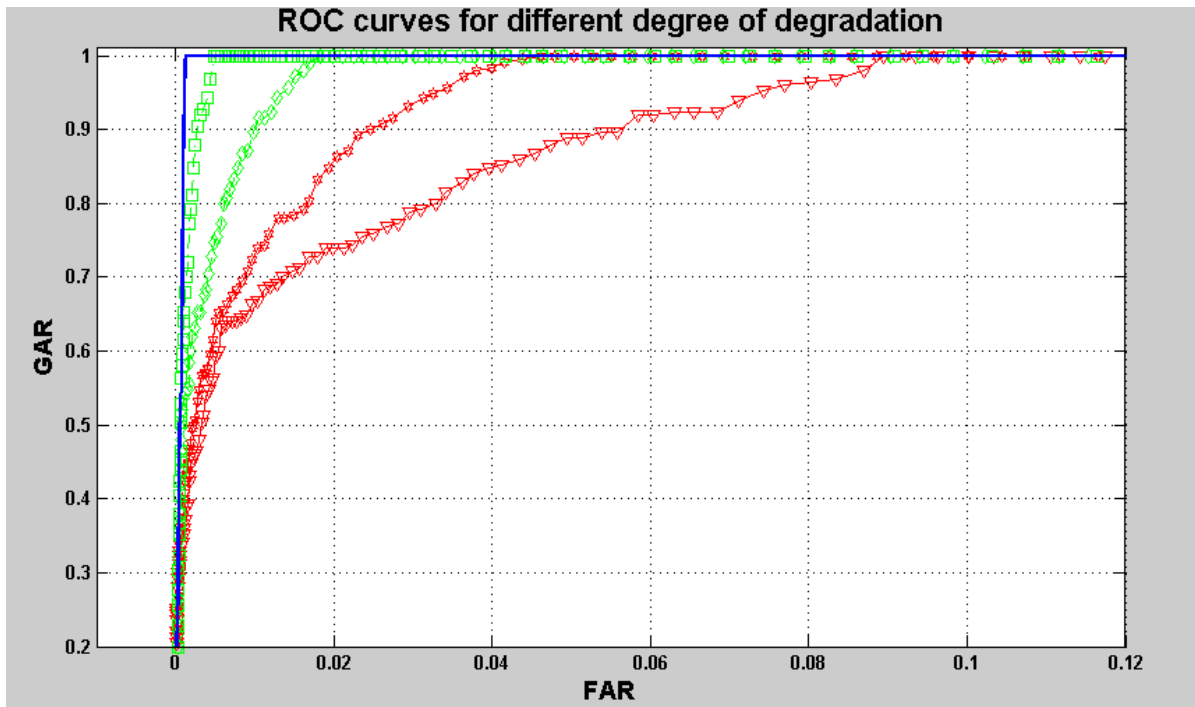


**Figure 6.1: Distance between the CC of 250 WM with pair-wise comparison.**

Next, we explore the robustness of the CC generation with different means of synthetic image degradation and respectively the identification of the watermarks using CCD metric. Synthetic distortion of the watermarks images was used to generate testing sample images. Table 6.1 displays the results of watermarks matching test (degree of similarity) based on the comparison of their CCD with four levels of imposed degradation. The degradation was with random synthetic distortion. The degradation was design not to exceed the indicated number of randomly changed CC value. The changes of CC's values depend on the number of relative location change in consecutive neighboring pixels. The limiting threshold level for number of changes in % value of the length of the CC is given in the first column of Table 6.1. The results from tests are given using the ROC values for the area under the curve (AUC) calculated at a given equal error rate (EER) and threshold.

**Table 6.1: EER, AUC, and the threshold for different degree of distortion.**

Random distortion Limit level	EER	AUC	Threshold
10%	.0025	.9989	231.9
15%	.0181	.9917	222.3
20%	.0322	.9823	213.9
25%	.0722	.9795	202.4

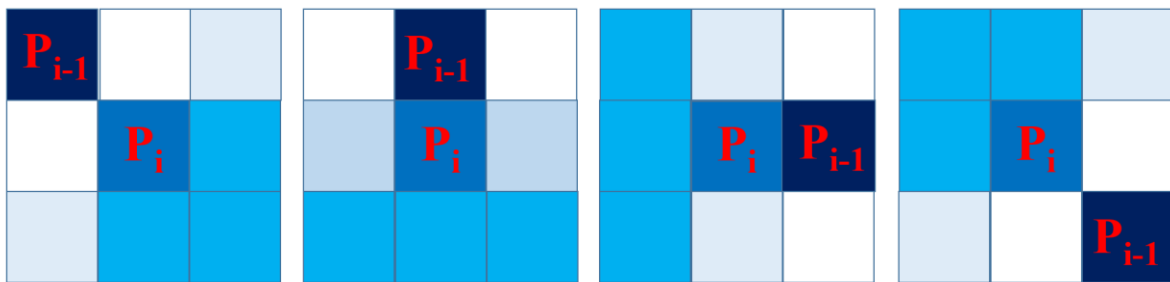


**Figure 6.2: ROC based on CCD for the degree of degradations given in Table 6.1.**

The effect of the synthetic degradation is noticeable but the AUC value continue to be close to one, indicating the robustness of the CCD metric. Please note that only the initial portion of the ROC is displayed (FAR from 0 to 0.12). The CC representation of the watermarks forms a high fidelity feature vector with high entropy and noise immunity, which even after a

substantial degradation continues to perform well. While the CC's values have been changed, it continues to retain a differentiating capability representing the watermark for comparison with other watermarks.

The synthetic degradation is intended to simulate the overall performance with noisy CCs. The CCs are generated after skeletonization of the binary watermark images. Most pixels in the skeleton, in general, have two neighboring pixels. The CC generation algorithm is designed to process the image starting from outside contours and going clockwise. This information is used to specify the directional change of the CC due to simulated digitization noise. Figure 6.3 displays four out of eight possible scenarios.



**Figure 6.3: The direction probability of CC formation. Due to the skeletonization, at pixel  $P_i$ , coming from pixel  $P_{i-1}$ , the direction of moving forward with higher probability is indicated with darker blue.**

The randomly selected index for value change  $P_i$  (dark blue), is examined backward to identify the neighbor  $P_{i-1}$  it follows (the darkest blue color). Based on the direction from which the CC is formed, the probability of the forward going direction is indicated with the degree of blue color.

#### **6.4 Alternative methods for watermark identification**

As already described, there are many methods available for shape and contour description. Most of them continue to be developed and applied to specific and different application. This thesis described a new method for description of watermark with CC after an initial preprocessing of the watermark images. In this chapter, modern methods for image matching are examined in order to compare their performance with the matching based on the new CCD metric.

The first chapters of this thesis reviewed the Content Based Image Retrieval (CBIR) process that is mostly applied to the description search and retrieval of watermarks. Featuring an exponentially increasing number of images on the web, the era of big data demanded the need for scalable systems which allow efficient indexing, retrieval, re-ranking, and browsing: The Bag-of-Words (BoW) model was introduced and viewed as the state of the art in Content-Based Image Retrieval (CBIR) systems since early 2000. In recent years, image representation based on the Convolutional Neural Network (CNN) has attracted more attention in image retrieval, and demonstrated impressive performance. Given this time of rapid evolution, this thesis explore the possibility of using CNN and BoW for watermark image retrieval methods.

According to [249], the feature extraction and quantization schemes can be used to classify current methods into three types, i.e., SIFT-based, one-pass CNN-based, and multi-pass CNN-based. The authors of the survey review milestones in BoW image retrieval, compares previous works that fall into different BoW steps, and shows that SIFT and CNN share common characteristics that can be incorporated in the BoW model.

In the next section of this chapter we present the results from applying CNN and BoW to watermark images.

### **6.4.1 Machine learning and image matching**

Recent advances in deep learning with applications in image matching and retrieval have prompted the test of NN for watermark images recognition. Several new approaches in the architecture, activation function, and training of NNs for deep learning have produced remarkable results [241, 243]. The new technology components include:

- Partially labeled data;
- Rectifier linear unit;
- Convolutional Neural network;
- Dropout neurons.

Most learning (training) algorithms are supervised or unsupervised. Partially labeled data provides a mixture of the two and provides additional opportunity. The Rectifier linear unit is becoming the standard activation function for the hidden layers of the deep learning NNs. In addition, the Linear or Softmax activation functions are used for the output layer. The Neuron dropout is a regularization technique in deep learning that prevents overfitting. This is important especially in multilayer, multiple neuron NN, in which a single or several dropout layers are included. One particularly appealing approach is to use a Convolutional neural networks (CNN). They have shown some record breaking performances with real image data and there is no need to generate features from the images in order to train/test the neural net. CNN input accept directly the image(s). The application of a five layer CNN for the recognition of watermark images is presented next.

### **6.4.2 Convolutional neural networks (CNN)**

After reviewing literature and implementing CNN structure in Matlab, with a limited exploration of the many optional settings, the NN structure that generated very good results is reported. Two convolutional layers with kernel size  $5 \times 5$  are used. The output map for the first was 8 and for the second 16. After each convolutional layer, scaling down with factor of 2 was used. The computational complexity could be indicated by performing a single epoch during the very first run of the training. Different batch size were explored to speed up the training. Several important notes follow.

It is critical to use image size that divides evenly by 4 (for the application of  $5 \times 5$  convolution kernel) and by 2 for each pooling layer of the NN. The possible starting image size depends on the convolutional kernel size as well. The size of the original watermark images in the Bernstein collection is  $100 \times 100$  pixels (divisible by 4 and applicable for use with  $5 \times 5$  convolutional kernel). Each layer of the implemented architecture changes the size of the input image to a map, one for each of the neurons (kernels / filters / features) in the layer. This image sizes after each layer are as follows. After the first convolution the  $100 \times 100$  size becomes  $96 \times 96$  pixels. There would be as many maps as the number of neurons in the layer. After the first down-scaling, the images are  $48 \times 48$  pixels. After the second convolution, they are  $44 \times 44$  which is an even number and could be scaled down by 2. In CNN terminology, the down-scaling is a result of the pooling, which in this case was an averaging over  $2 \times 2$  pixel areas.

The CNN are extensively used in the area of “Deep Learning” and are a subject of intensive research. New developments are improving the performance of new architectures. In many applications, pre-trained layer-wise network could be used on line (cloud based). The examples

presented in this thesis are with supervised learning using the classical sigmoid activation function. Recent advancements in deep learning appear to use different activation functions rectified linear unit (ReLU), parametric ReLU, exponential linear unit (ELU), soft exponential, soft plus, bent identity and others [241, 242]. Deep learning is considered the third stage of the NN development, a concept that was initiated in the 1940s. In 2006 Hinton proposed a radical new way of training deep NN. Recent advances in high performance GPU allows the use of larger and deeper architectures NN [243].

### **6.4.3 Watermarks dataset for use with CNN**

In general, the watermark dataset has a single representative of a given watermark class or, at most, several representatives which are grouped in a subclass. To train the NN the need for a large dataset is eminent. The solution to this problem is data augmentation. It is achieved by applying “label preserving transformations”. Several augmentation techniques have been developed specifically to test the performance of algorithm in the area of digital processing of images of historical documents. They include Kanungo noise, character degradation and geometric distortion [244].

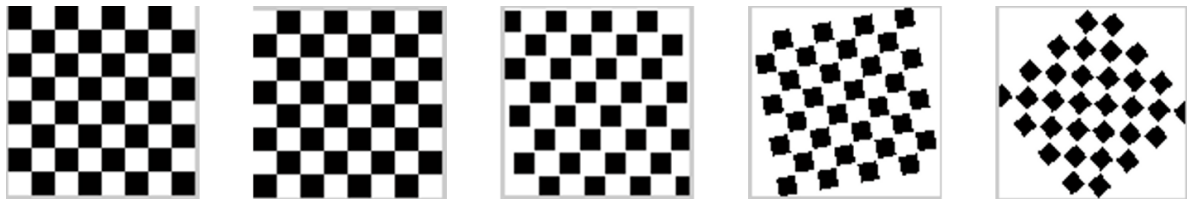
In [245], Kanungo et al. introduced a model to distort character edges. First, the value of some foreground and background pixels are changed (foreground pixel changes to background and vice versa) according to a probability function based on selected pixel distances. The process continues by using a non-linear local selection process and a morphological closing operation to smooth the characters edges. This model works well with binary images and is widely used to simulate the presence of noise for assessing the performances of different document analysis methods.



The character degradation model proposed in [246], which is specifically designed to mimic typical noise regions in ancient documents such as ink splotches and white specks or streaks due to the age of the manuscripts. This kind of noise mostly appears in the neighborhood of the characters and can lead to touching and broken characters.

The geometric surface deformation was proposed by J. Liang et al. in [247] to evaluate their restoration algorithm. First, this model maps every point in the plane of the original image to a point on a curved surface. The curved surface is then divided into planar strips (e.g. quadrilaterals) that are mapped back to the image plane. In this thesis work, similar degradation model with different modulation surfaces such as sinusoidal and parabolic is used.

To generate multiple representative watermark images, every image in the original set was synthetically transformed with ten different methods. The result was a set with sufficient size to be used for testing with CNN. Perhaps there are other transformations that could be used. Also, some transforms are not included intentionally. For example, the translation transformation does not result in any useful (different) result because the moved (translated on the plane) images are the same as the original. The shearing and stretching is easily made with *tformarray()* function of Matlab and the results are similar to the use of transformation using *fitgeotrans()* function. To better visual representation, Figure 6.4 displays the effect of the used transforms on a checkerboard image.

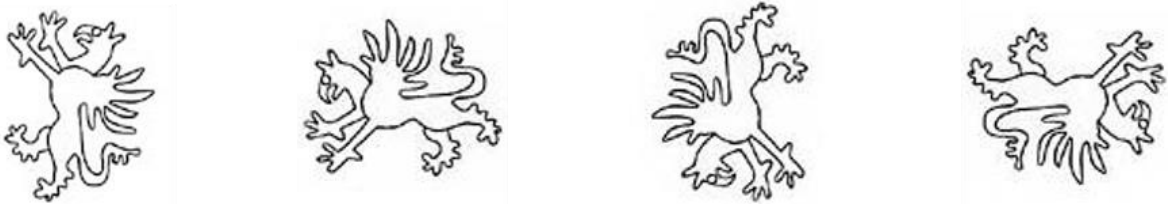




**Figure 6.4: Synthetic image distortions. From left to right top row: original checkerboard image, rotation, affine, reflective and non-reflective similarity. Bottom row: projective, polynomial, piece-wise linear, barrel and pin-cushion transforms.**

The used transforms are listed next with watermark image examples and short explanations. First, a rotation was performed consecutively three times at 90 degrees counter-clockwise to generate three additional images from the original watermark image.

### Rotation



In the next examples of transformations, the first of the images is the original watermark image. The following four are the results of the transforms performed on the rotated original images.

**Affine transformation** – the *affine2d()* and *imwarp()* functions of Matlab are used. The particular transformational matrix T is:

$$T = \begin{bmatrix} sc & -ss & 0 \\ ss & sc & 0 \\ tx & ty & 1 \end{bmatrix}; \quad \text{where } sc = \text{scale} * \cos(\text{angle}), \text{ } ss = \text{scale} * \sin(\text{angle}), \text{ } tx = 0 \text{ and } ty = 0.$$



**Reflective similarity** - The particular transformational matrix T is:

$$T = \begin{bmatrix} sc & -ss & 0 \\ a*ss & a*sc & 0 \\ tx & ty & 1 \end{bmatrix}; \quad \text{where the parameter } a = -1.$$



**Non-reflective similarity** - The particular transformational matrix T is:

$$T = \begin{bmatrix} sc & -ss & 0 \\ a*ss & a*sc & 0 \\ tx & ty & 1 \end{bmatrix}; \quad \text{where the parameter } a = 1 \text{ or is simply omitted.}$$



**Projective** - A good illustration of the result is with a checker board above. The particular transformational matrix T for this example is:

$$T = \begin{bmatrix} 1.0 & 0 & 0.001 \\ 0.1 & 1.0 & 0.001 \\ 0 & 0 & 1.0 \end{bmatrix};$$



**Polynomial** – Using reshaping with random points, the image warping is performed. The polynomial transform was performed with *fitgeotrans()* function of Matlab.



**Piece-wise linear** – The *fitgeotrans(movingPoints, fixedPoints, 'pwl')* is used with  $\text{movingPoints} = [10\ 10; 10\ 30; 30\ 30; 30\ 10]$  and  $\text{fixedPoints} = [10\ 10; 10\ 30; 32\ 32; 30\ 10];$



**Locally-varying with sinusoid** – A mesh-grid  $[x_i, y_i]$  of the image is created and two new vectors  $u$  and  $v$  are formed  $u = x_i + a_1 \cdot \sin(\pi \cdot x_i / \text{imid});$  and  $v = y_i - a_2 \cdot \sin(\pi \cdot y_i / \text{imid});$  After that the new image is created using *makeresampler()* and *tformarray()* functions of Matlab. The *tformarray()* function is like *imtransform()*, but is intended for problems of mixed input/output dimensionality and provides greater user control or customization.



The next figure helps to illustrate the barrel and pin-cushion distortion in images which are typically associated with the lenses of the optical system.

**Radial barrel distortion** – The transformation is performed with conversion from Cartesian to polar coordinates and back by varying the amplitude of the cubic term for the radius and consecutively using the *tformarray()* function of Matlab. For this specific example, the radius is calculated using  $s = r + a * r^3$  with  $a = 0.0001$  value.



**Radial pin cushion distortion** – The pin-cushion distortion is achieved by using a negative value for  $a$ . In this particular example,  $a = -0.000005$ .



As indicated previously, the size of the input images are important for the proper operation of the CNN. The design of the neural network could be made to operate with different size input images. This would require addition of sublayers for image auto-resizing. It is better and more efficient to use an image set with the appropriate size for the CNN design. Several CNN

configurations were used. For example, configurations c8/p2/c16/p2 and c6/p2/c12/p2 which were explained previously.

All of the transforms listed above have several parameters that can be modified to achieve different degree of change. The parameter values for the particular examples are given. One critical thing is that the transforms change the size of the original images. To resolve the problem, the original 100 x 100 binary images are padded with zeros: 100 x 50 on the left, 100 x 49 on the right, 50 x 199 on top, and 49 x 199 on the bottom. This results in a 199 x 199 (odd) size images with a defined horizontal and vertical center lines. After the transformations, the resultant images are cropped back to 100 x 100 size. This way they are ready for use with a CNN without implemented layers for auto-resizing.

The following Figure 6.5 displays an example of 10 original watermark images. Figure 6.6 presents some of the transformed images.



**Figure 6.5: WM images used in CNN example.**

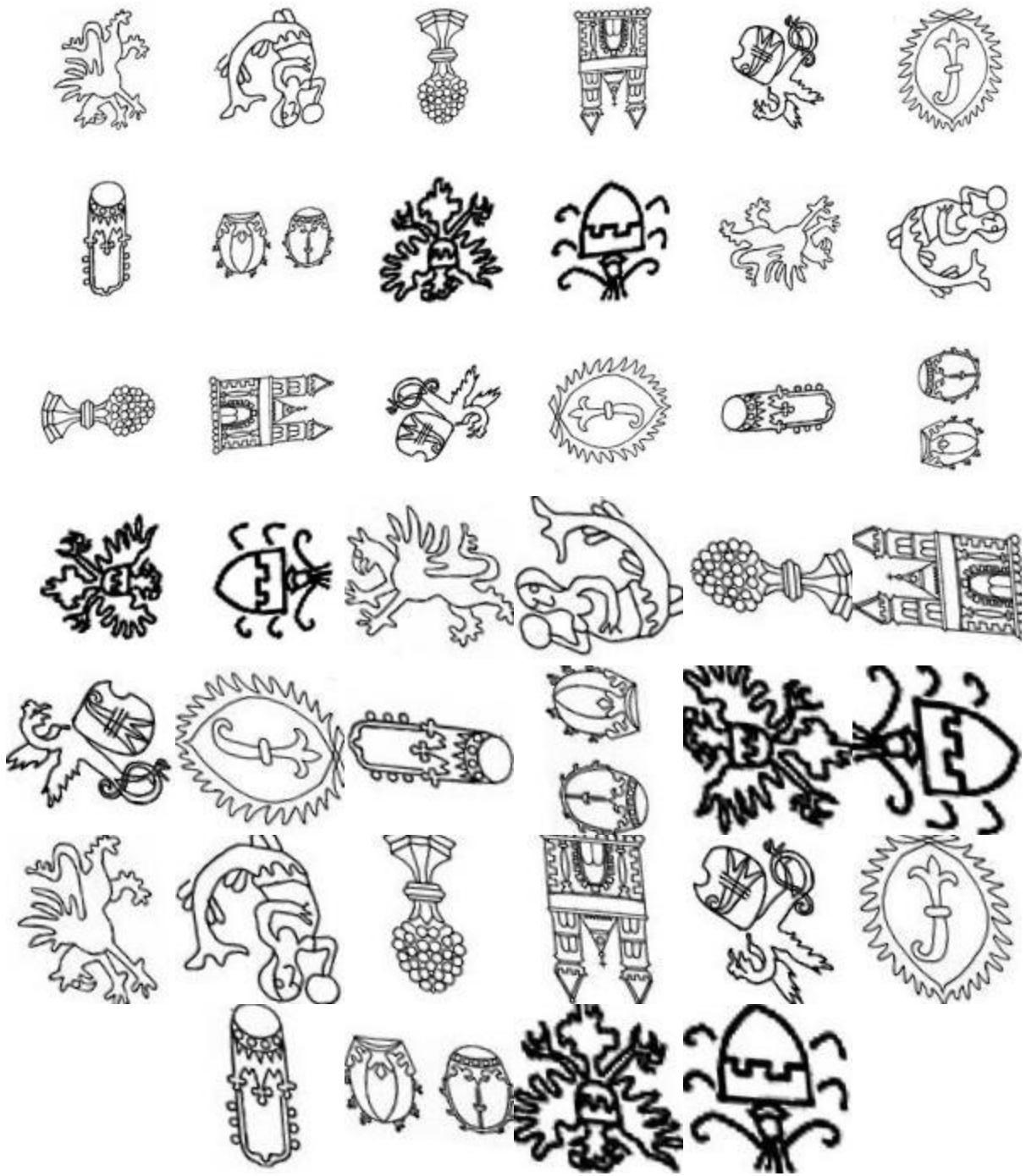


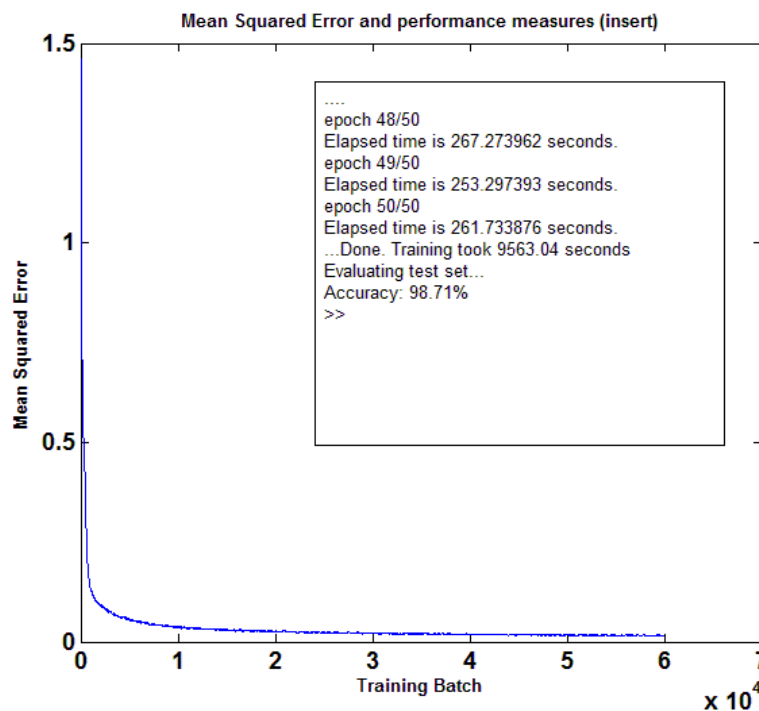
Figure 6.6: Examples of transformed images for CNN training.

All original images of the watermarks and the transformed images are displayed in Appendix B. The transformed and scaled images are displayed in Appendix C.

#### 6.4.4 Experimental results of watermarks matching using CNN

In the Matlab implementation, the CNN's architecture is stored in the structure with a cell array containing each of the five feature extraction layers. The output layer parameters are stored in a separate object. The CNN has six layers: an input layer, 2 convolution layers, 2 subsampling (or "pooling") layers, and an output layer which performs the classification. It took some time to figure out many aspects of the CNN structure definition and training/test set generation. The CNN was trained with the augmented images and tested with the original images.

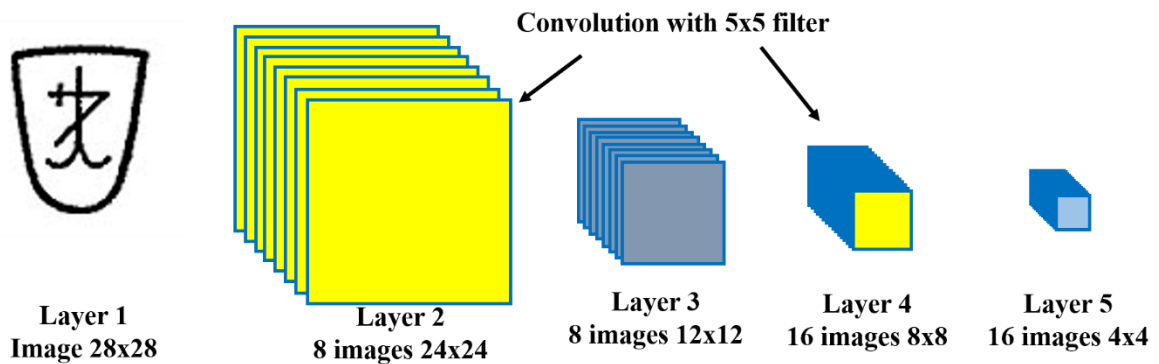
The performance of the CNN with 100 x 100 size input images and 50 epochs is given in Figure 6.7.



**Figure 6.7: CNN train with 50 epochs and performance test results (insert) for 100 x 100 size images.**



Working with the original image size produced good results but the individual epoch execution time was relatively long. After exploring with the original size of the images, the image size was reduced in order to speed up the training and respective testing. The next experiment was performed with images scaled down to one half on each axis (one quarter of the original image area). The original images were scaled to 56 x 56 pixels with MATLAB function *imresize(I, 0.5)*. After each NN layer, the size changes from layer (1) 56 x 56, to 52 x 52 pixels at layer (2), after the first convolution with 5 x 5. The image size at layer (3) is 26 x 26, after the first pooling with 2 x 2 mask, and reduces further to 22 x 22, after the second convolution with 5 x 5 at layer (4). The final size is at layer (5) is 11 x 11 after the second pooling.



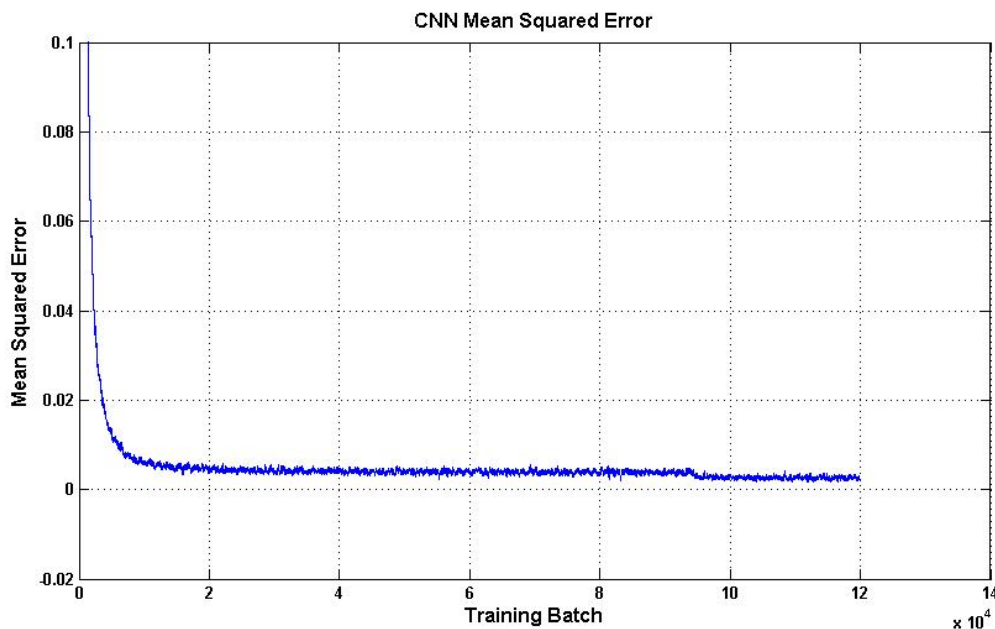
**Figure 6.8: An example of c8/p2/c16/p2 CNN architecture.**

Running the CNN implementation with several different layer configurations results in different times per epoch and accuracy performance. Following the resizing explanation, the CNN configuration can be written as c8/p2/c16/p2. Figure 6.8 illustrates the architecture of the CNN. The 8 neurons of the first convolutional layer generate 8 separate maps and respectively, the 16 neurons of the second convolutional layer generate 16 separate maps. Using smaller number of neurons executes faster but achieves lower accuracy. For example, c8/p2/c16/p2 architecture increased the running time per epoch (370 sec) without improving the accuracy

substantially compared to c6/p2/c12/p2 configuration (~280 sec) on an AMD A8-4500M, 1.8 GHz Laptop to complete a single epoch.

To further explore the performance and potentially improve the speed of execution, the use of images scaled to 28 x 28 pixels were used. The smaller image size did not degrade the performance result. The Matlab function *imresize(I 0.275)* was used in order to get 28 x 28 pixels size images from the original images with size 100 x 100 pixels. Again, the size changes from layer (1) 28 x 28, to 24 x 24 pixels at layer (2), after the first convolution with 5 x 5. The image size at layer (3) is 12 x 12, after the first 2 x 2 pooling, and reduces further to 8 x 8, after the second convolution with 5 x 5 at layer (4). The final size is at layer (5) is 4 x 4 after the second pooling.

Figure 6.9 presents the mean square error convergence for 100 epochs of training.



**Figure 6.9: Mean Square Error for c8/p2/c16/p2 CNN.**

The expanded dataset of 2500 samples were divided in training, validation and testing sets with 7:2:1 ratios using stratified partitioning. With 50 epochs of training, the achieved rank-1 accuracy was 98.71%. The accuracy increased to 99.25% with 100 epochs of training. This is

a very impressive result which demonstrates the capabilities of CNN regardless the relatively small training sample size. It took about 230 seconds on AMD A8-4500M 1.8 GHz laptop to complete a single epoch. Scaling down the CNN architecture speeds up the training process but also results in lower accuracy. CNN configured as c6/p2/c12/p2 executed a single epoch on average for about 170 seconds with accuracy decreased below 90%.

#### **6.4.5 Bag of visual words**

In this section, to demonstrate the possible use of Deep Learning, Bag of Visual Words (BoVW) method for classification of digital images of paper watermarks is demonstrated. At least theoretically, the method is applicable for identifying to which class a given watermark belongs based on the constructed BoVW using the features of the watermarks of this class.

One approach is to use Bag of Visual Words (BoVW) for classification of WM images. The original applications of this method was develop as bag-of-words (BoW). BoW has multiple implementations based on the features and classifiers used. The images however do not have words that are explicitly available, and the BoW model cannot be applied directly. Fortunately, using a method for feature extraction, the words of the image can be generated. The Scale-Invariant Feature Transform (SIFT) [248] is attractive to use because it is comprised of feature detector and descriptor, which is invariant to image translation, scaling, and rotation, and is robust to local geometric distortion. SIFT is also is partially invariant to illumination changes but it is not of interest to the watermarks binary images. With SIFT, an image can be transformed into a collection of local feature vectors, which can be viewed as prototypes of words in text. For this thesis, following several references and code examples, several BoVW were considered. In many of the references, authors use SIFT at several scales, spatial

histograms, Fisher vectors, spatial pyramids for extracting features. K-means is usually used for visual word dictionary construction. SVM is often use for classification. In this thesis, SIFT was used for features generation and Elkan's k-means for fast dictionary construction. Implementation of fast k-means algorithm was not critical for this example with a small data set but it is critical for large datasets.

The steps for implementing the BoVW can be summarized as follows:

1. Extract features.
2. Learn “visual vocabulary”.
3. Quantize features using visual vocabulary.
4. Represent images by frequencies of “visual words”.

In general, the watermark data sets have single representatives for a given watermark. Obviously, we cannot train the BoVW with a single image. As already discussed in the application of CNN, the solution to this problem is data augmentation. It is achieved by applying “label preserving transformations”. The same methodology described in “Deep Learning with CNN” section was used to generate 10 sets of 40 images with 100x100 pixels each.

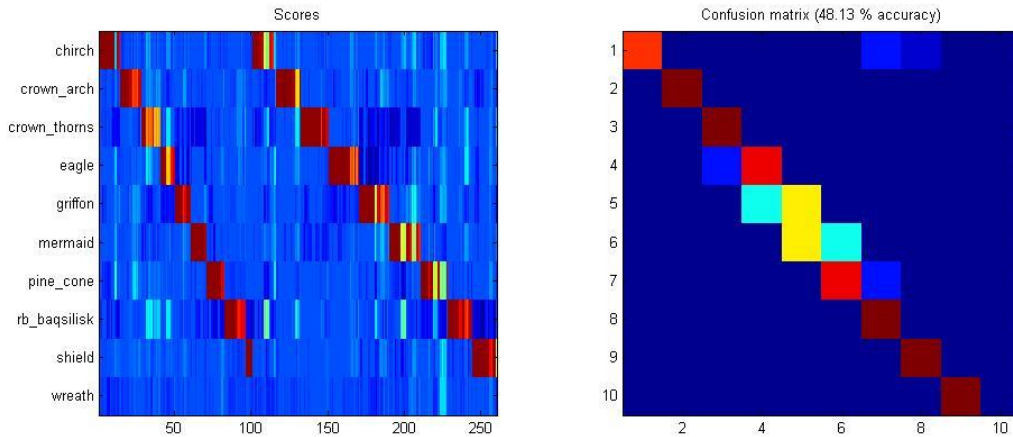
Each augmented set of watermarks was stored in a separate subdirectory, forming a class for training and testing. Disjoint sets of images were randomly assigned for training and testing and a limit of 300 words was used. Several methods for features extraction (generation) were tested before SIFT was finally used due to its advantages described above.

The purpose of the k-means clustering was to partition the set of vectors into groups around common mean vector, so the clustering process finds, in principle, the best dictionary or codebook to *vector quantize* the data. The Lloyd-based k-means algorithm was used,

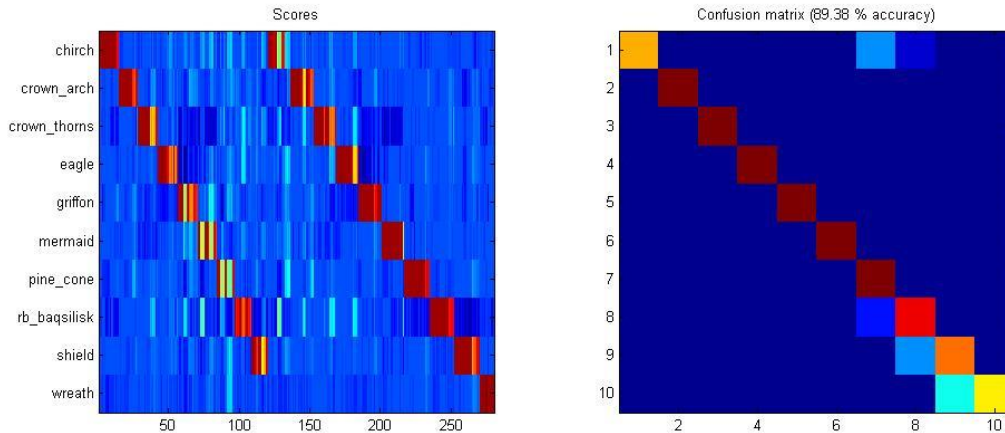
alternating between optimizing the cluster centers and the data-to-center assignments. The process generates a matrix containing the cluster centers and a vector with assignments of the input data to the clusters. The cluster centers are also called *means* because it can be shown that, when the clustering is optimal, the centers are the means of the corresponding data points.

#### 6.4.6 Experimental results using BoVW

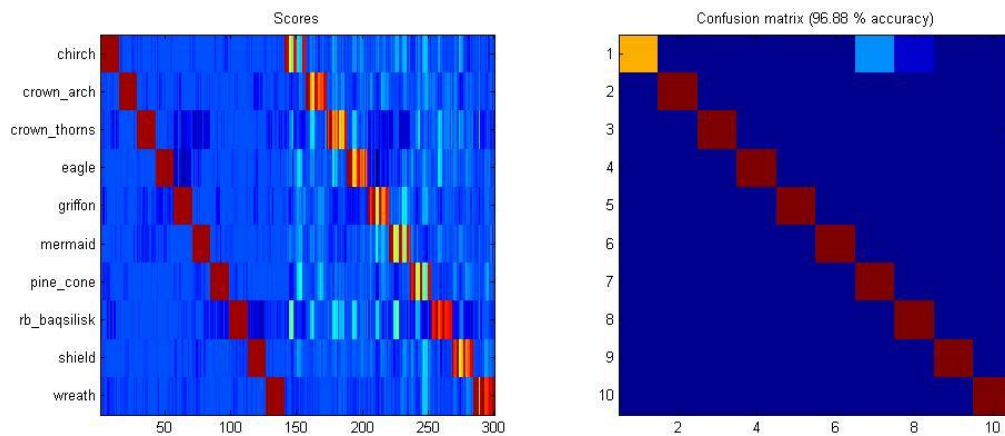
Figure 6.10 presents the classification scores and the confusion matrices with achieved accuracies for three different training sessions. Experiments with different training size generated different results and, as expected, the classification improves with the increased number of training images. With only 15 training images per class, the achieved accuracy very poor even for the limited set of only 10 classes Figure 6.10.



**Figure 6.10: Classification accuracy of 48.13 % for 10 watermarks using limited training with 150 randomly selected watermark images.**



**Figure 6.11: Classification accuracy of 89.38 % with 30 images per class selected randomly for training (300 out of 400) and a disjointed set of 10 images for testing.**



**Figure 6.12: Classification accuracy of 96.88 % with 39 images per class selected randomly for training (390 out of 400) and the rest for testing.**

The heat-maps in Figure 8 present the scores (top) and the confusion matrix (bottom). A rank-1 classification accuracy of 89.38% was obtained when 30 randomly selected images per class have been used for training. When 39 out of 40 images per class were randomly selected for training and the remaining single image for testing, a maximum rank-1 classification accuracy of 96.88% was achieved.

It is possible to explain the high accuracy with the very small data set. On the other hand, the watermark images do not carry reach textual information. The result is a very good conformation about the applicability of BoVW for watermark images classification. The

reported results has to be viewed as exploratory work and prove of concept, rather than as development of a persistent system for watermark classification with BoVW.

## **6.5 Conclusions**

Chain codes are an effective descriptor of watermarks with high accuracy and sufficient differentiation capability. With the developed method for numerical comparison of CCs with different length, the similarity measurement of CCs provides fast matching of watermarks.

The CC representation of the watermarks forms a high fidelity feature vector with high entropy and noise immunity, which even after a substantial degradation continues to perform well. While the CC's values have been changed, it continues to retain a differentiating capability representing the watermark for comparison with other watermarks. Testing the effect of the synthetic degradation indicates noticeable changes of the CCD but the AUC value continue to be close to one, indicating the robustness of the new distance metric.

The use of CNN has the advantage of direct image comparison after initial training. One problem with the application to watermarks is the deficiency of large datasets for training. To illustrate the possibility of watermarks recognition with “deep learning”, a large dataset was built for training and testing using synthetic image degradation. The reported test results indicate the applicability of the CNN for watermark image recognition.

## CHAPTER 7

### WATERMARK COMPOSITE FEATURE VECTOR

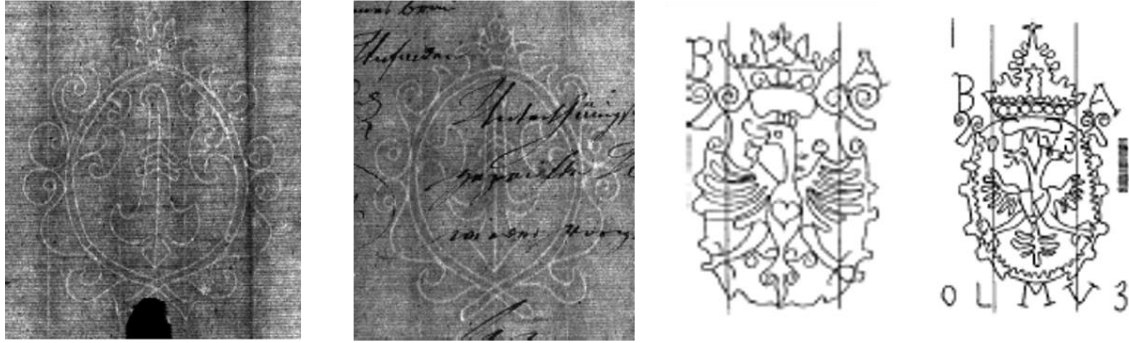
#### **7.1 Introduction**

There are many examples when watermarks belong to a group of very similar design and vary by elements added to, removed, or changed from the original design. It is desirable to be able to describe these minor differences and account for them during the identification process. We address the need for a complete description of the details of the watermarks by including additional descriptors to form what we named a Composite Feature Vector.

#### **7.2 Watermarks complexity**

Thousands of watermarks have been designed and used through the centuries. Thousands are reproduced and digitized. Watermarks vary from very simple and easy to replicate to very complex watermarks which were designed with the intent to be difficult to copy and prevent forgery of the marked paper. In this thesis, a framework is presented for automatic identification of watermarks based on the generated chain codes. It is important to note that more complex watermarks offer more content that could be used for identification based on modern image processing algorithms. Some watermarks require rigorous preprocessing before construction of good skeleton suitable for CC generation. Many watermark images are of good quality but the watermarks are with complex structures. Figure 7.1 displays some examples of such watermarks.





**Figure 7.1: Examples of watermark images requiring rigorous preprocessing or with complex motives.**

For images of watermarks with high complexity, standard image based recognition will work better based on the reach texture content. Images that result in good quality binary representation and moderate complexity of motives could be represented with extended chain code arrays.

### **7.3 Extended chain code arrays**

The recording of the extended chain code arrays is based on the iterative processing of the skeleton image of watermark and removal of the contour layers from outside of the watermark towards the center. The result is a set of CCs which are grouped and recorded with predefined order. The proposed format for recording is to use tags structure elements analogous to the tags used in HTML documents. The reason is flexibility and ability for on line posting and processing. HTML uses brackets < > to encase the tags. For instance, a paragraph is considered to be an html element constructed with the tags <P> and </P>. In the same fashion, the proposed chain code element (extended CC arrays) is defined to start with <CC> and end with </CC> tags.

#### 7.4 Structure of the Composite feature vector

As previously described, the CC consist of the  $x$  and  $y$  indexes of the CC start pixel and a sequence of coded direction of the neighboring pixels in clockwise traversing of a contour. As previously demonstrated, some of the watermarks elements are not included in the CC generated from the first pass of the algorithm. During the first pass, the algorithm only generates a CC for the outside boundary contour. The remaining watermark elements can be detected and described after consecutive removal of the described portions of the watermark. This is achieved using the re-generative property of the CC and subtraction from the original watermark binary image. This process is repeated until a predefined number of passes is completed or until the generated set of CC describes completely the watermark elements. For example, a criteria could be based on the rule of “no pixels left” or “less than a threshold number pixels left” after series of CC generation, re-generation and subtraction from the original image”. Alternative solution is to limit the number of iterative levels of description. As shown, the CCs are able to capture high degree of characteristic features which are sufficient differentiating representation of complex contours.

Recording the coordinates of the CC’s start pixel position allows exact reverse transform from CC to the image. In addition, the two-value set of the start pixel coordinates is expanded with a third number which is the length of the CC. Borrowing the tag notation from HTML, as described previously, each CC record in the CFV is formatted in the following way:

<CC>

<SHA> 03, S, H, A</SHA>

<01,X,Y,L1> X1, Y1, L1, c1, c2, c3..., cL1</01,X,Y,L1>

<02,X,Y,L2> X2, Y2, L2, c1, c2, c3..., cL2</02,X,Y,L2>

<03,X,Y,L03> X3, Y3, L3, c1, c2, c3..., cL3</03,X,Y,L03>

<MD>d1, d2,dc3..., dN</MD>

<MD>d1, d2,dc3..., dN</MD>

<!-- This could be any comment or additional information -->

</CC>

Following the <CC> header tag which indicated the start of a CFV, the <SHA> indicates the enclosed information is for the watermark CFV number of CC, length of the short and long sides of the MBB, and the angle of rotation of the MBB. The information is represented with a three chain codes. “S” indicate the length in pixels of the width (short side) and “H” indicates the height (longer side) of the MBB. “A” is the angle of rotation to align the MBB with the x-y coordinates expressed floating point in the range 0.0 to 359.9. The top level information ends with </SHA>.

Please note that the real records do not have spaces following the comma separated values. The notation here is presented with spaces for better readability.

Next, the tag is <01,X,Y,L1> X1, Y1, L1, c1, c2, c3..., cL1</01,X,Y,L1> and includes start of the first CC array with location coordinates of the start pixel (X1, Y1) and the length of the CC (L1). The </XYL 01> indicates the end of the first CC array. In the same way, the <02,X,Y,L2> indicates that the second CC has start pixel at location (X2, Y2) and length of L2 decimal characters. The </02,X,Y,L02> indicates the end of the second CC array. This continues and the number of CC arrays has to be as indicated in the top level information which in the example above is three CC arrays.

An expansion of the CFV includes additional data. The meta data is included in the <MD>d1, d2,dc3..., dN</MD> section of the CC. If present, it may indicate different ratios

that could be are computed and used as additional discriminating parameters. For more complex watermarks, values associated with the convex hull, centroid, and others could be added to the CFV. The Anker watermark used in the examples has an area of 1715 pixels, convex hull area of 3,566 pixels, and MBB area of 4928 pixels. This values could be used directly or to calculate useful ratios to further differentiate watermarks with close similarity. The HTML language uses <meta .....> for meta data tag format. This can be freely used in the CFV in addition to the specifically included meta data. Also, as typically used in html files, comment tags may be included in <!-- --> bracketed format.

### **7.5 Composition and use of the CFV**

The proposed format facilitate the parsing of the CFV string. In the example, <SHA> tag indicates the top level information (header information) about the CFV.

In general, CCs have different length. Including the length of the CC gives a second advantage: the database of the CC records can be indexed and divided in subsections by the length of the CC. For example, CC with length in the range 600 to 700 could be grouped together in a search sub-folder. Search for matching and identification will be performed based on the CC length and will speed up the process especially for large databases.

The proposed tags for the recording of the CFV follows the established notation of HTML. It is designed to provide an easy and unambiguous way for parsing the information. The scalability of the CC's structure provides flexibility for description of even very complex watermarks (multiple contour shells). The structure could be customized for a particular dataset (database). Following the proposed structure, will provide standardization and interoperability.

In general, the first CC will have the longest length – it is the outmost contour. In some cases, it is possible that the watermark has a motif that is in the upper left corner and will be processed first by the algorithm. To resolve these cases, the CC generation algorithm records the multiple chain codes sorted by length in descending order.

## 7.6 Examples with CFV application

The assembling of the content of the composite feature vector (CFV) is demonstrated with two examples cases. First, we present the CFV for relatively simple watermark. Second, we demonstrate the differences in the CFVs for a watermark from the same family which includes additional wires, resulting in additional markings on the paper. Structurally, the CFV for the two scenarios will be as follows:

### Simple contour watermark

<CC>

<01, S, H, A>

<01, X1, Y1, L1>c1, c2, c3..., cL1</01, X1, Y1, L1>

</CC>

### Watermark with additional motive

<CC>

<02, S, H, A>

<01, X1, Y1, L1>c1, c2, c3..., cL1</01, X1, Y1, L1>

<02, X2, Y2, L2>c1, c2, c3..., cL2</02, X2, Y2, L2>

</CC>

Following the <CC> header of the CC string, for the simple watermark the <01, S, H, A> indicates the watermark is represented with a single CC. “S” and “H” indicate the length in

pixels of the width (short side) and height (longer side) of the MBB. “A” is the angle of rotation to align the MBB with the x-y coordinates expressed floating point in the range 0.0 to 359.9. The <01, X1, Y1, L1> indicates that this single CC has start pixel at location X1, Y1 and length of L1 decimal characters.

Following the <CC> header of the more complex watermark, the <02, S, H, A> indicates the watermark is represented with two CCs. The <01, X1, Y1, L1> indicates that this first CC has start pixel at location (X1, Y1) and length of L1 decimal characters. The <02, X2, Y2, L2> indicates that this second CC has start pixel at location (X2, Y2) and length of L2 decimal characters.

Figure 7.2 displays a group of class *Bell* watermarks. Starting with the simple watermark more elements have been added. The watermarks are used to illustrate the generation of the CFV.



**Figure 7.2: Watermarks of class *Bell*. From simple figure to one with two additional elements.**

Using the algorithm for generation of the chain code, the images are processed iteratively until less than 50 foreground pixels remain in the image after consecutively removing the regenerated image sections. Figure 7.3 displays the intermittent image results for watermark 0069005. The algorithm generates the CC for the outermost contour, regenerate the processed



22211010000000000012233232323232323232376676767676767677667734444444444444444  
5455666666666666666666666666666634432322222222222223434444444444444445222 </01,33,70,493>  
<02,5,36,106>0777777777770111111211104555655555555677777777623333333  
33333454555655555511111121110123333333333334</02,5,36,106>  
</CC>

**Watermark with additional motive 006009004:**

<CC>  
<02,84,85,1.9>  
<01,6,46,238>122221221212212111111101110101010010001000000000000  
0007000700700707077777777767676767676667666666666666665666566566565  
65565565555545554545454445444444444444444444344344343433433343233333  
3232323223222322222222222222</01,6,46,238>  
<02,30,70,203>0122112121212212212212221222222222222222222322232221007666  
666667011222232221226770001007666656566766656666667666766667766667676677  
6767766704444444444444445566666656676544323222222222344444444443444</02,3  
0,70,203>  
</CC>

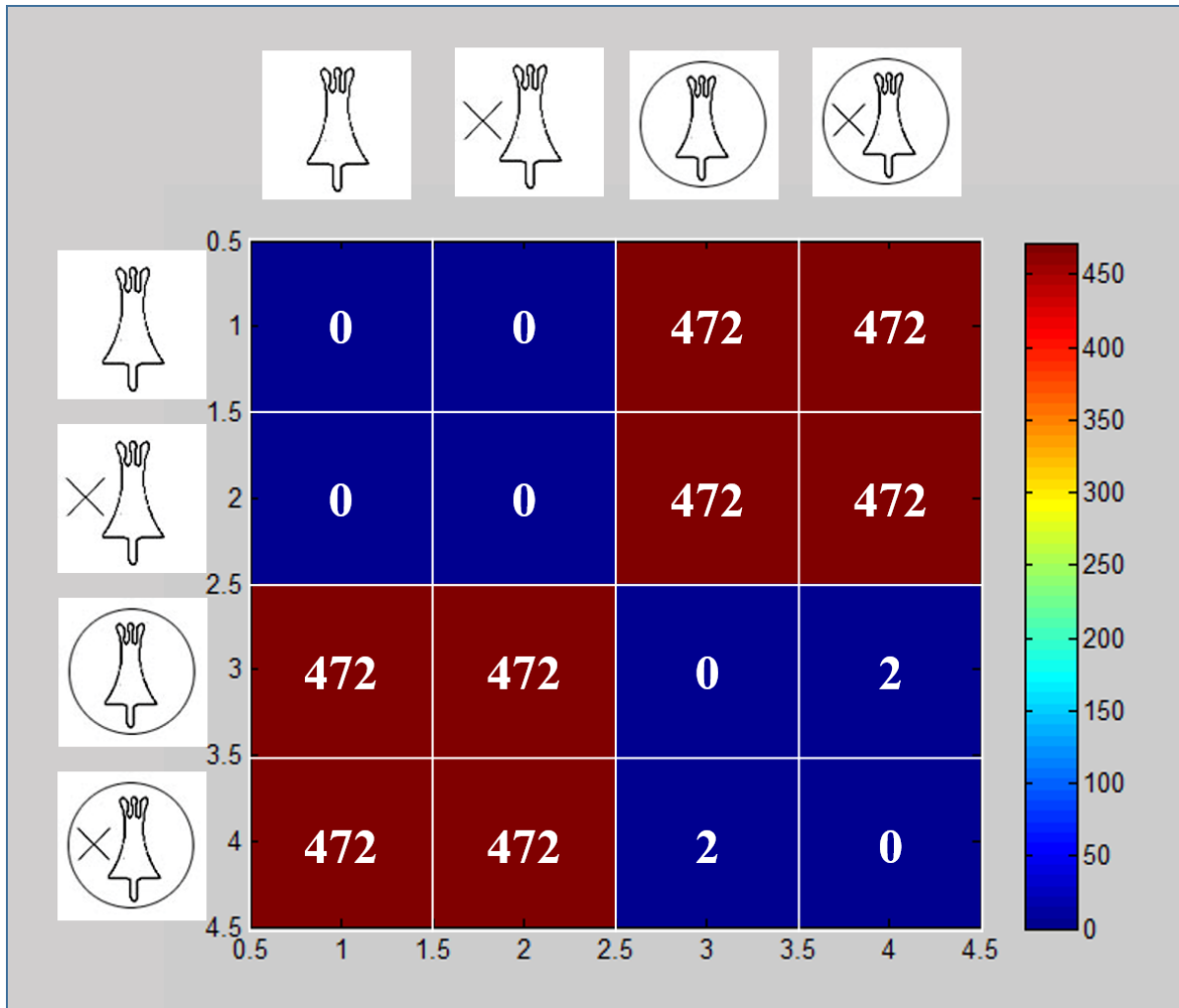
**Watermark with additional motive 006009005:**

<CC>  
<03,84,85,1.9>  
<01,6,46,238>122221221221212111111101110101010010001000000000000  
0007000700700707077777777767676767676667666666666666665666566566565  
65565565555545554545454445444444444444444444344344343433433343233333  
3232323223222322222222222222</01,6,46,238>  
<02,33,70,222>012211212121221221221222122222222222222222322232221007666  
666667011222232221226776666666700123222222100766665665676665666666766  
676667666676766767767677667044444444444444455666666666665443232222222234  
4444444443444</02,33,70,222>  
<03,13,37,88>77777777711111111126555555555667777777762333333333355555  
555556211111112233333333333</03,13,37,88>  
</CC>

Figure 7.4 illustrates the use of the CFV for watermarks recognition. The images in Figure 7.2 all have the *Bell* watermark and three of them have additional elements. As can be seen, based on the current approach, the matching is for images 006009001 and 006009003 and also for images 006009004 and 006009005. It is possible that the CFV generation has to be modified in order to capture special cases such as this one with an additional enclosing circle. One possible option is for the user to be able to interactively request RST processing after the



removal of the outermost contour. In this case, the matching will be based on the Bell shape and the watermarks will be grouped as a class.



**Figure 7.4: Recognition of similar watermarks with additional motives using CFV.**

### 7.7 Conclusions

Chain codes demonstrate capability to be used as descriptors of watermarks. The RST-invariant CCs combined with the CCD provide new opportunity for deterministic representation and matching of watermarks. The variability and complexity of watermarks requires the extension of the representation with a composite feature vector (CFV).

The structure of the vector provides scalability to allow complete description even of the most complex watermarks. The natural extension with meta-data, comments, graphics, and other standard tags available HTML can be easily implemented. Standardization of the CFV structure is the base for accurate and fast comparison of the watermarks they represent.

The proposed tags for the recording of the CFV follows the established notation of HTML. It is designed to provide an easy and unambiguous way for parsing the information.

Including the length of the CC gives a second advantage: the database of the CC records can be indexed and divided in subsections by the length of the CC.

## CHAPTER 8

### SUMMARY OF PRESENTED WORK

#### **8.1 Discussion**

In this dissertation, I present a novel method for automatic classification, identification (matching) and retrieval of watermark images based on CC descriptors. The novel approach for generation of unique watermark descriptor based on RST-invariant CC is presented. In addition, the proposed method resolve the recto and verso variability of watermark images. This results in a RSTV-invariant representations of watermarks.

A new term is introduced to name the watermark signatures based on the chain codes and the composite feature vector C-grams.

The unique C-grams are truly reversible providing high-rate lossless compression.

The development of a novel distance measure for CC comparison is also presented. The performance of the algorithm on large datasets is demonstrated using watermark images from well-known library catalogue collections.

The existing techniques for watermark digitization are reviewed with respect to the produced image quality. The applicable methods for binary image formation, including morphological operators, are discussed and demonstrated.

The work is completed as a part of the CODICES Project, which is conducted by a multidisciplinary team at UMKC, including faculty and graduate students from English, History, Chemistry, and Computer Science and Engineering departments. I would like to thank again Dr. Reza Derakhshani, Dr. Jeffrey Rydberg-Cox, Dr. Virginia Blanton, Dr. Nathan A.

Oyler, and the graduate students Chainy J. Folsom, Melissa M. Morris, and Sara Derakhshani for their collaboration during the work on the CODICES project.

## **8.2 Conclusions**

One summary statement for the work presented in this thesis is the introduction of the C-gram as a deterministic descriptor of watermarks. With a C-gram each watermark has its unique ID.

We conclude that the work described in this thesis, titled “Framework for automatic identification of paper watermarks with chain codes” successfully achieved five important goals:

1. Generation of unique numerical watermark descriptor based on chain code.
2. New method for generation of rotation and scale invariant chain codes (RST-invariant CC).
3. Development of new method to compare two chain codes with different size – distance metric for CCs similarity measurements, CCD.
4. Exact re-generative, very high ratio lossless compression of watermark images based on chain code.
5. Development of a composite feature vector (CFV) for complete description and comparison of watermarks.

### **Work accomplishments related to paper watermarks:**

- The overall achievements of the work described in this thesis can be summarized with the introduction of the C-gram of a watermark.

- The overall novel contribution of the work presented in this thesis provides for the first time a deterministic method for generation of complete numerical description of watermarks.
- The CC length can be used for database indexing and accelerated searching.
- The scalability of the CFV provides method for complete description of more complex watermarks.
- The proposed structure of the CFV provides standardization and interoperability.
- An alternative, non-contact watermarks reproduction method is presented applicable to digitization of valuable documents in fragile state.

**Work accomplishments related to the general field of pattern recognition:**

- Development of method for generation of rotation and scale invariant chain codes: RST-invariant CC.
- Development of method for mirror image variance test and detection: RSTV-invariant CC.
- Complete description of complex contour (and shape boundaries).
- Exact re-generation of complex and nested contours.

**8.3 Future work**

This thesis present the framework for automatic identification of watermarks based on their description with chain codes. The concept has been introduced and the implementation of working models for all necessary steps have been demonstrated using 250 watermarks from the Bernstein collection (Appendix A). The initial test indicate 100% recognition accuracy on

the 250 watermark images even after degradation. Next natural thing is to expand the application of the proposed method to more datasets.

Presenting the framework to the scholars engaged in the field of filinogranology may result in many valuable suggestions which will result in improvements of the methodology.

The future work may be summarize in three main areas:

1. Testing with expanded watermark datasets.
2. Refinement of the CFV.
3. Development of an integrated GUI for watermark image processing.
4. Complimentary extension of the filigranology C-gram integrated work

environment with deep learning based capabilities for watermark image identification.

The first area will reveal deficiencies and prompt further development of the proposed method. Most likely the performance within single databases will be good. Some additional components may be needed to provide interoperability across data collections. The results of work in this direction will provide necessary input for work in areas 2 and 3.

It is possible to optimize the CCD. CC has a limited space presentation (8-character alphabet) which is *a priory* known or even can be selected. For example, negative and positive decimal values have been considered. Using alphabet  $\{-4, -3, -2, -1, 1, 2, 3, 4\}$  and absolute values for distance measure was explored but not completely and it is still an option. The location of the alphabet characters (values) is also a potential area for optimization.

In this thesis, F8 CC representation was used and explored to result in the introduction of watermarks C-grams. It may be beneficial to explore some the other chain codes and compare performance with the F8 chain codes.

The CFV is intentionally designed as a scalable representation. Main work in this area could be done on standardization of the metadata capturing.

Developing of GUI for watermarks preprocessing and CFV generation will be very beneficial to the scholars in the filigranology field. The application can be open source based on a general public license or on line accessible application hosted by an appropriate organization. My current vision for that is to design an interactive environment in which the user can navigate with selection of options to accomplish a complete and accurate CFV generation for any possible watermark. The use of such integrated environment will produce many new recommendations and valuable information about the possibilities and limitations of the proposed method for automatic watermarks identifications.

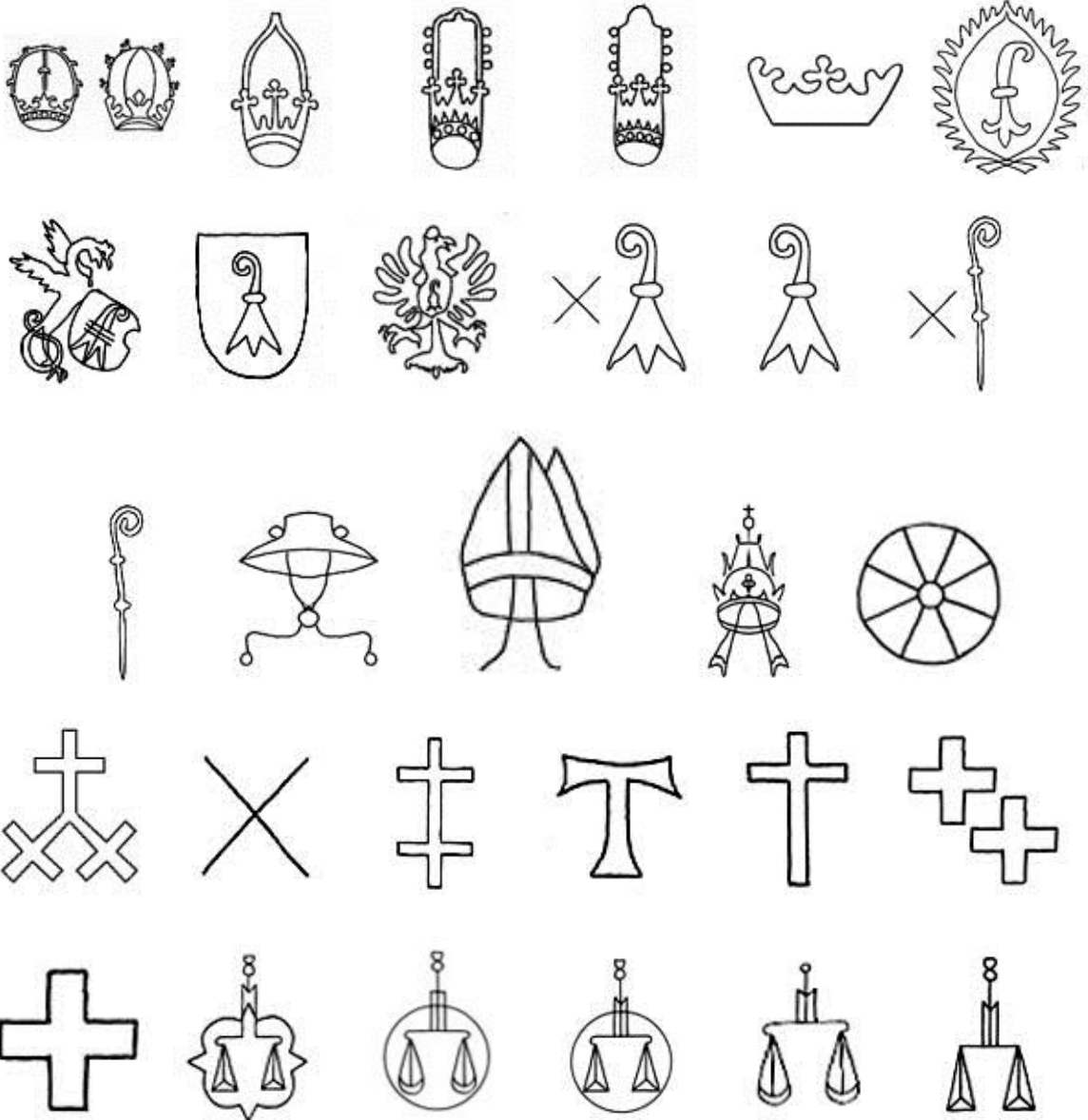
With the development of the new deep learning methods for image identification, the user should have the ability to select the alternative methods. This is even more relevant for watermarks with high complexity.

It is my desire and believe that the proposed framework for automatic identification of watermarks will benefit the work of many scholars working in the field of filigranology.

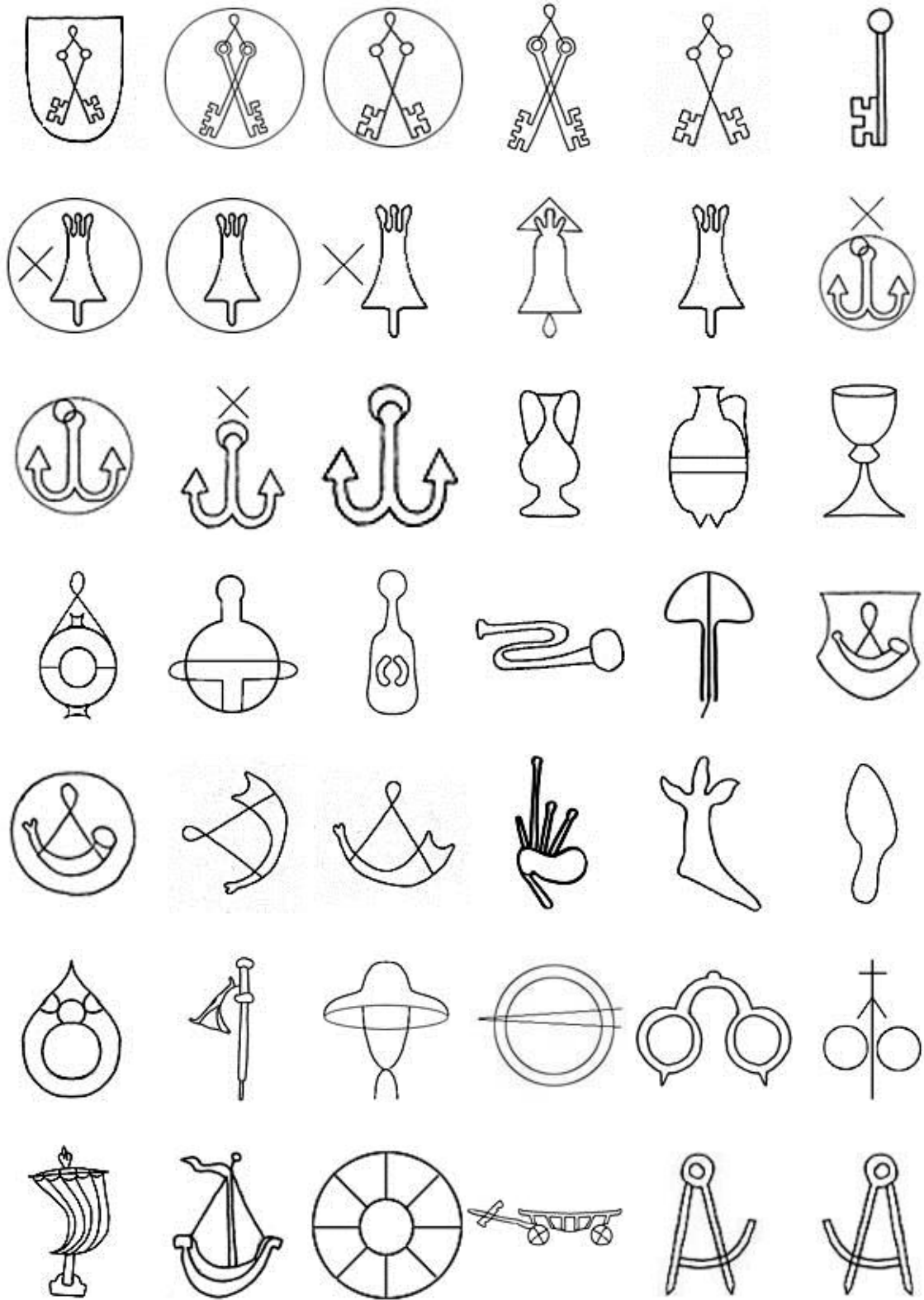
## APPENDIX A

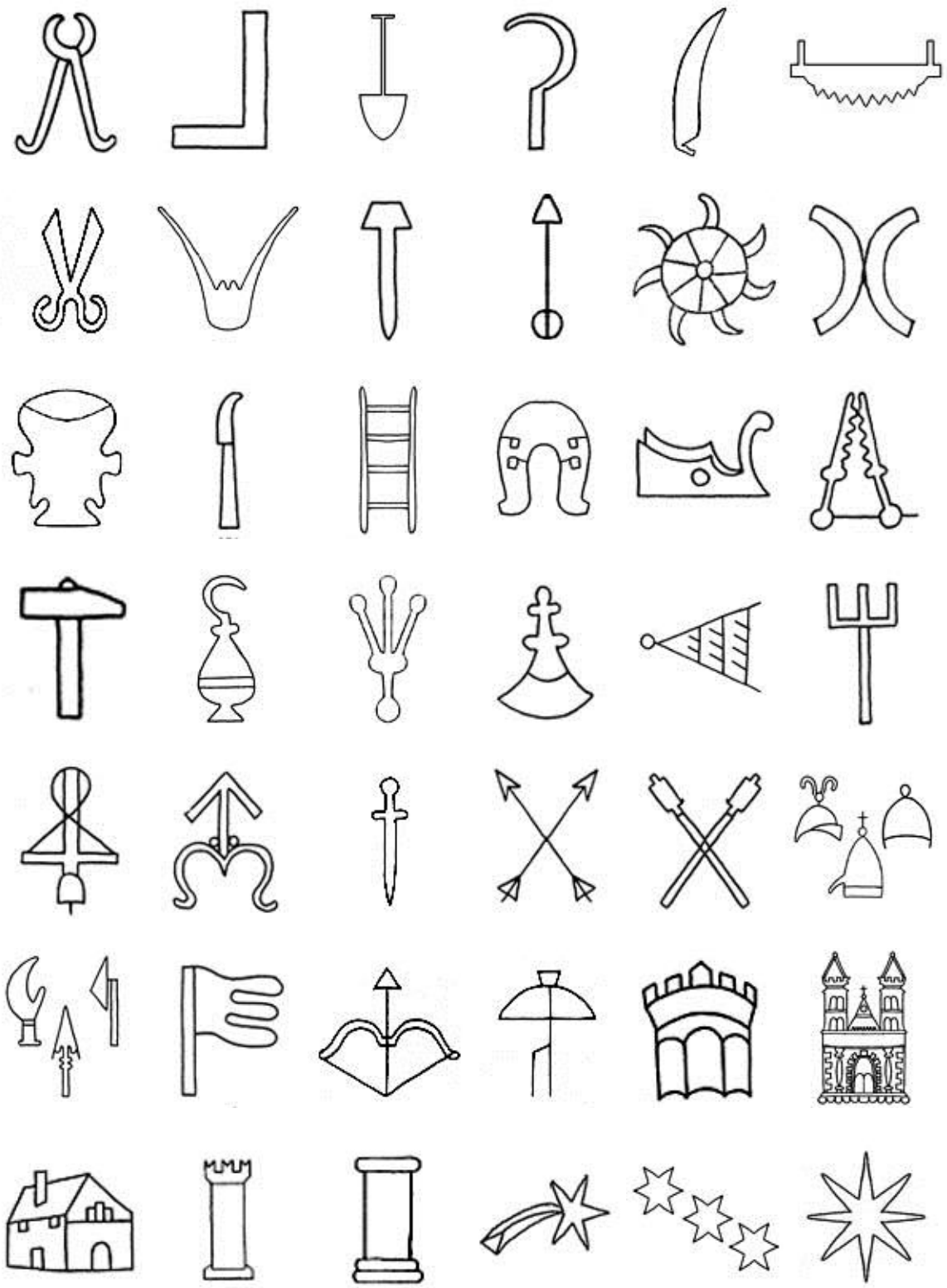
### WATERMARKS FROM THE BERNSTEIN COLLECTION USED IN THIS STUDY

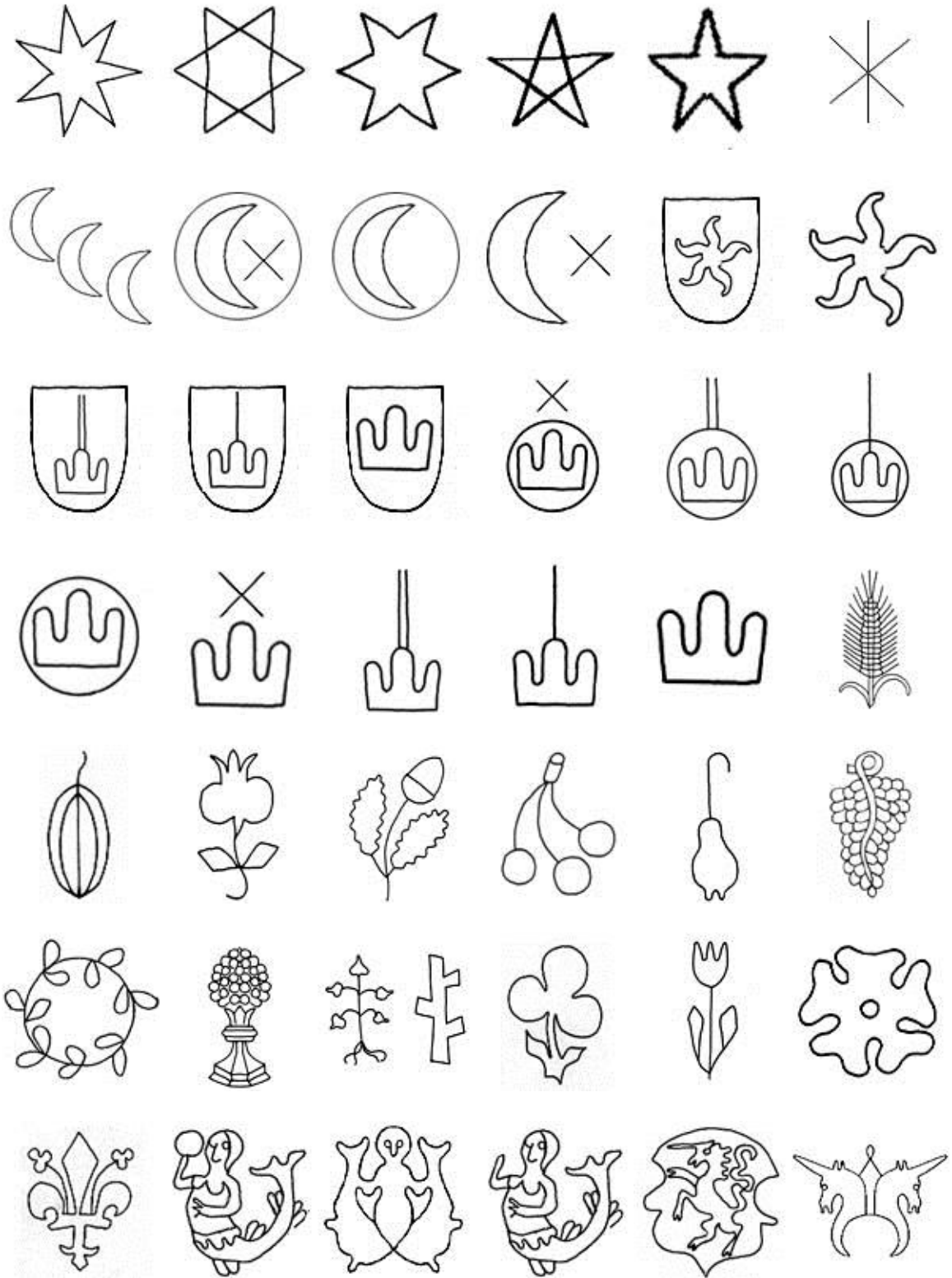
Two hundred and fifty images of watermarks from the Bernstein collection used in this study.

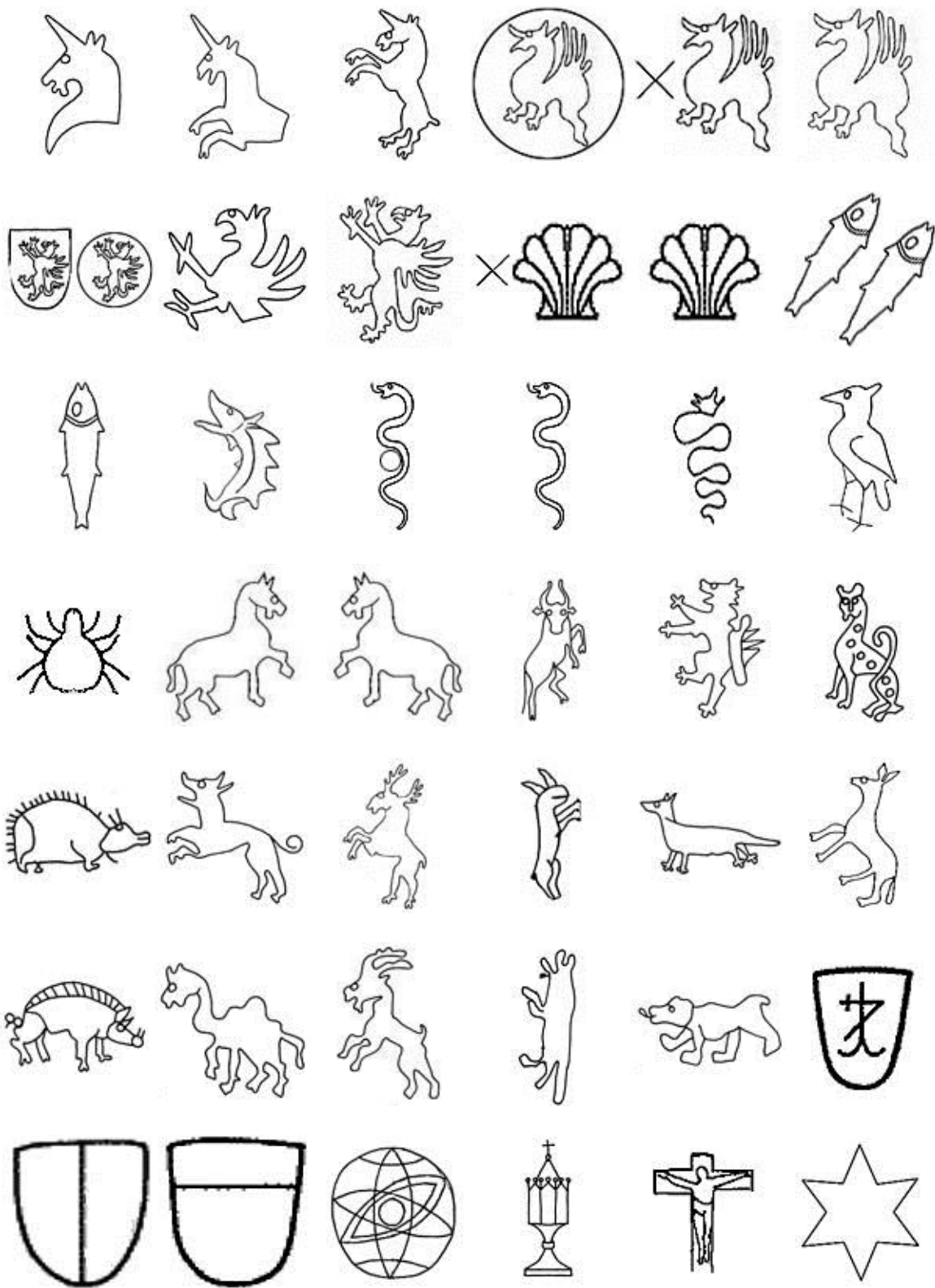


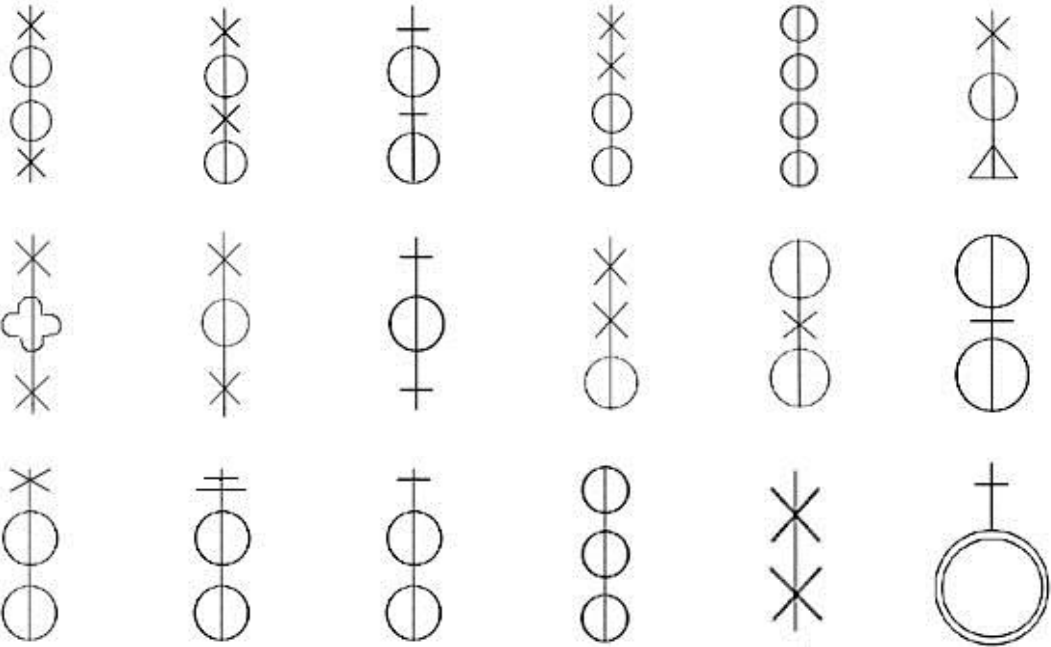
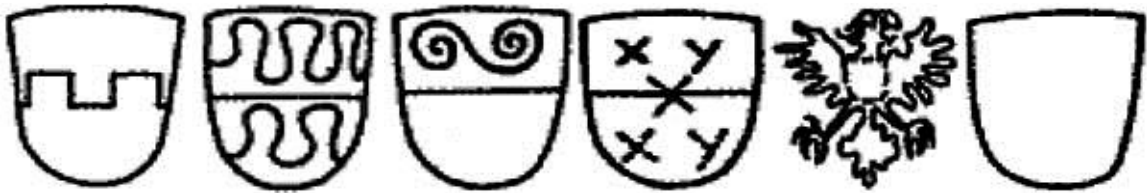
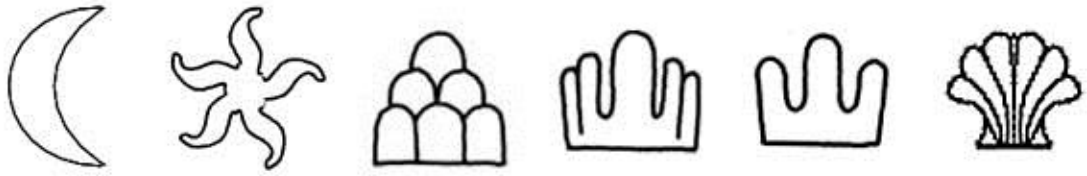


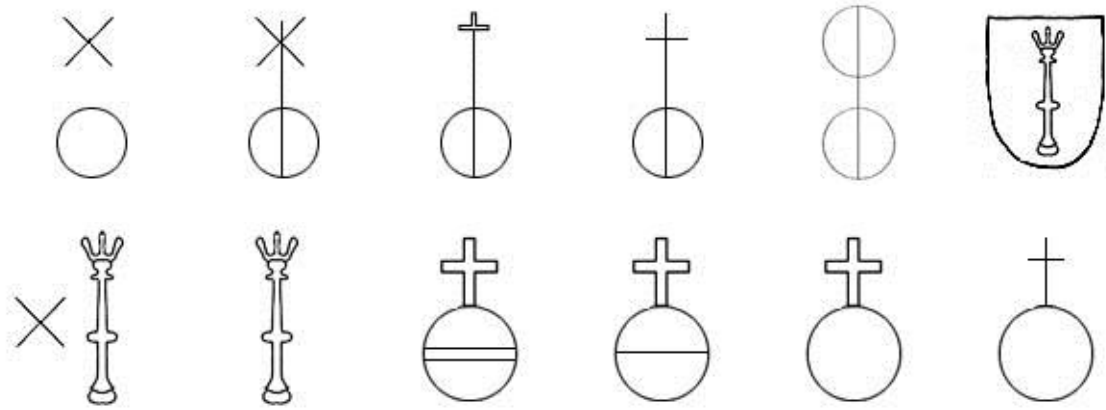








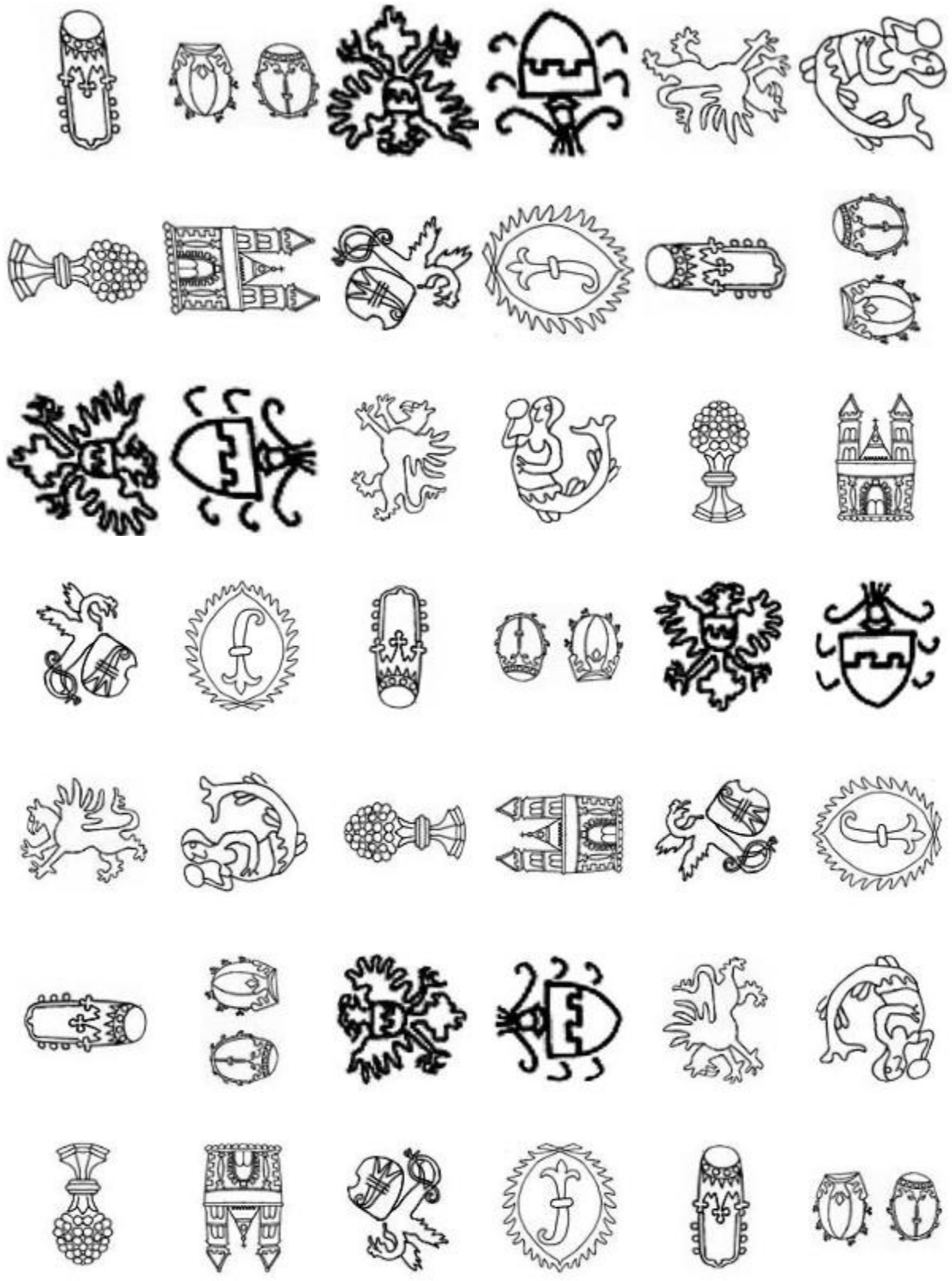




APPENDIX B

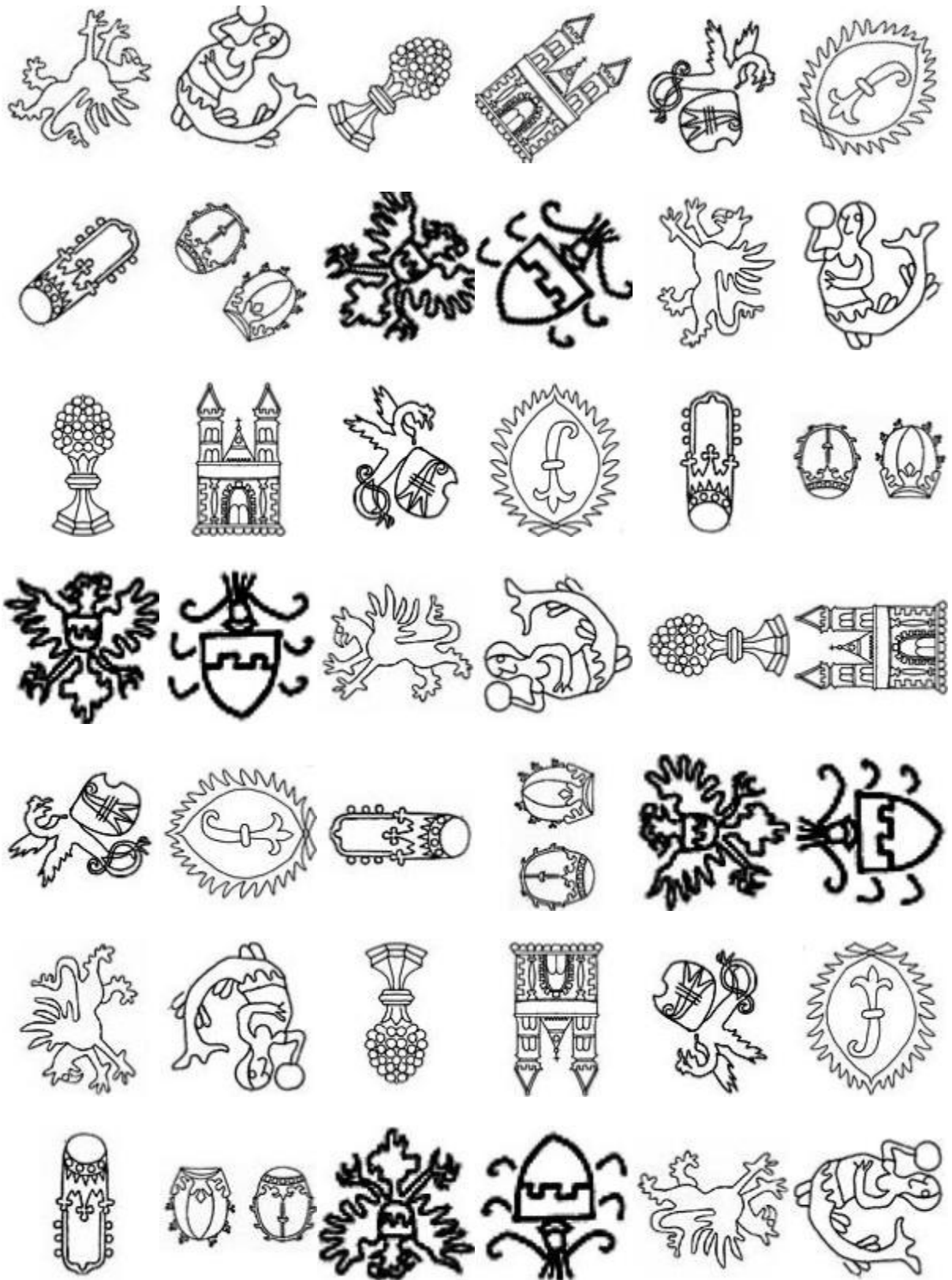
THE ORIGINAL AND TRANSFORMED IMAGES OF WATERMARKS

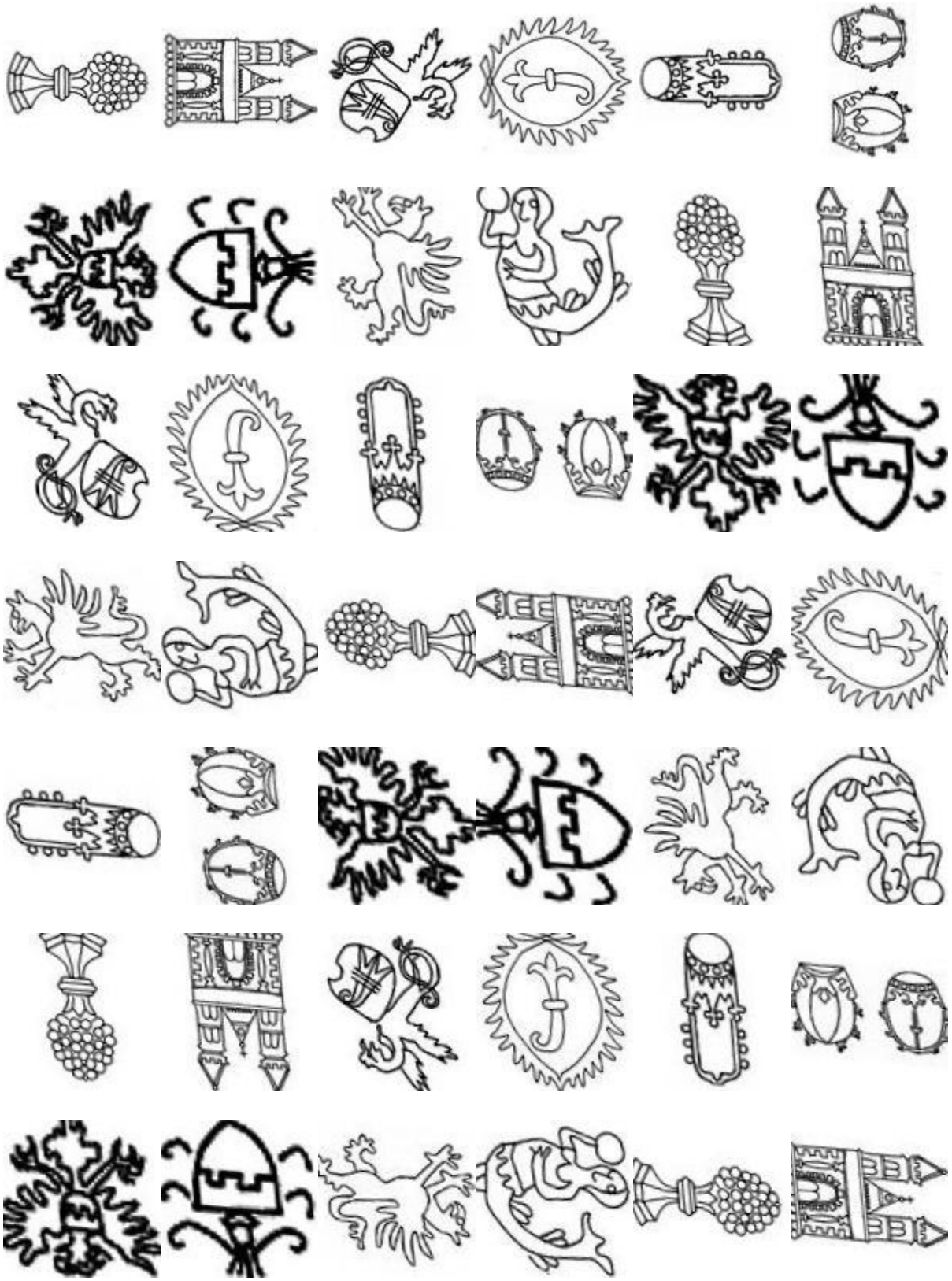


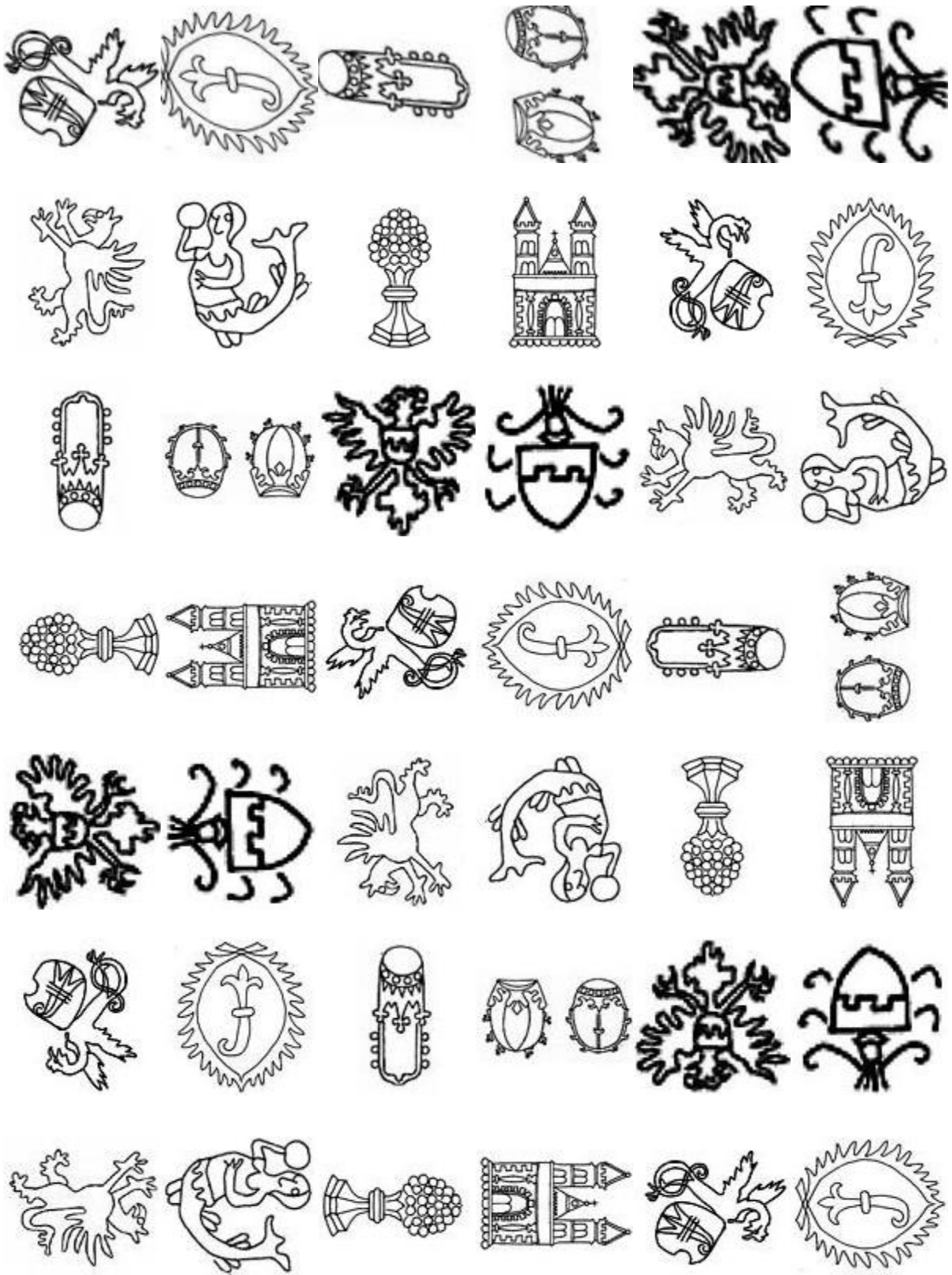


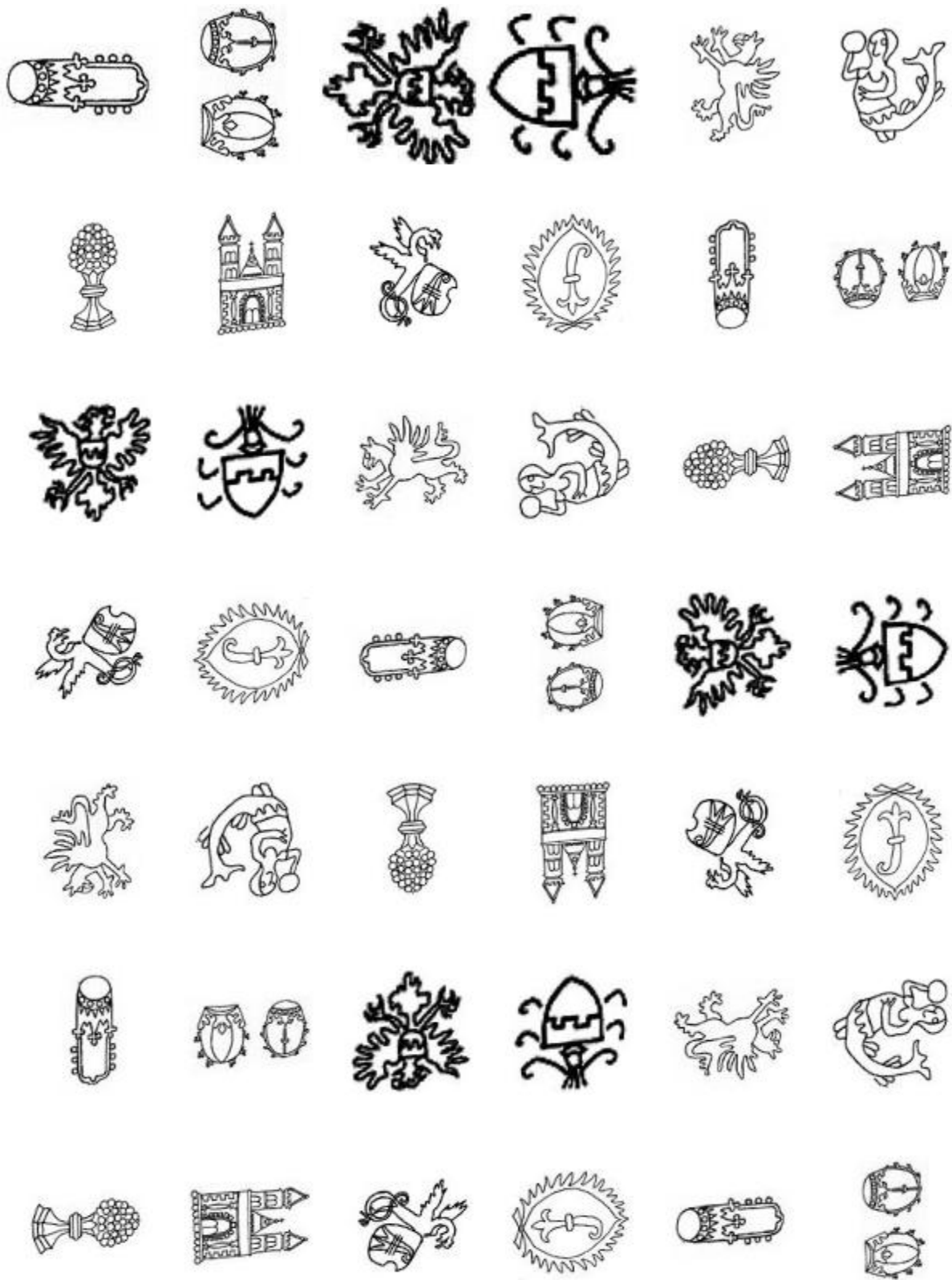






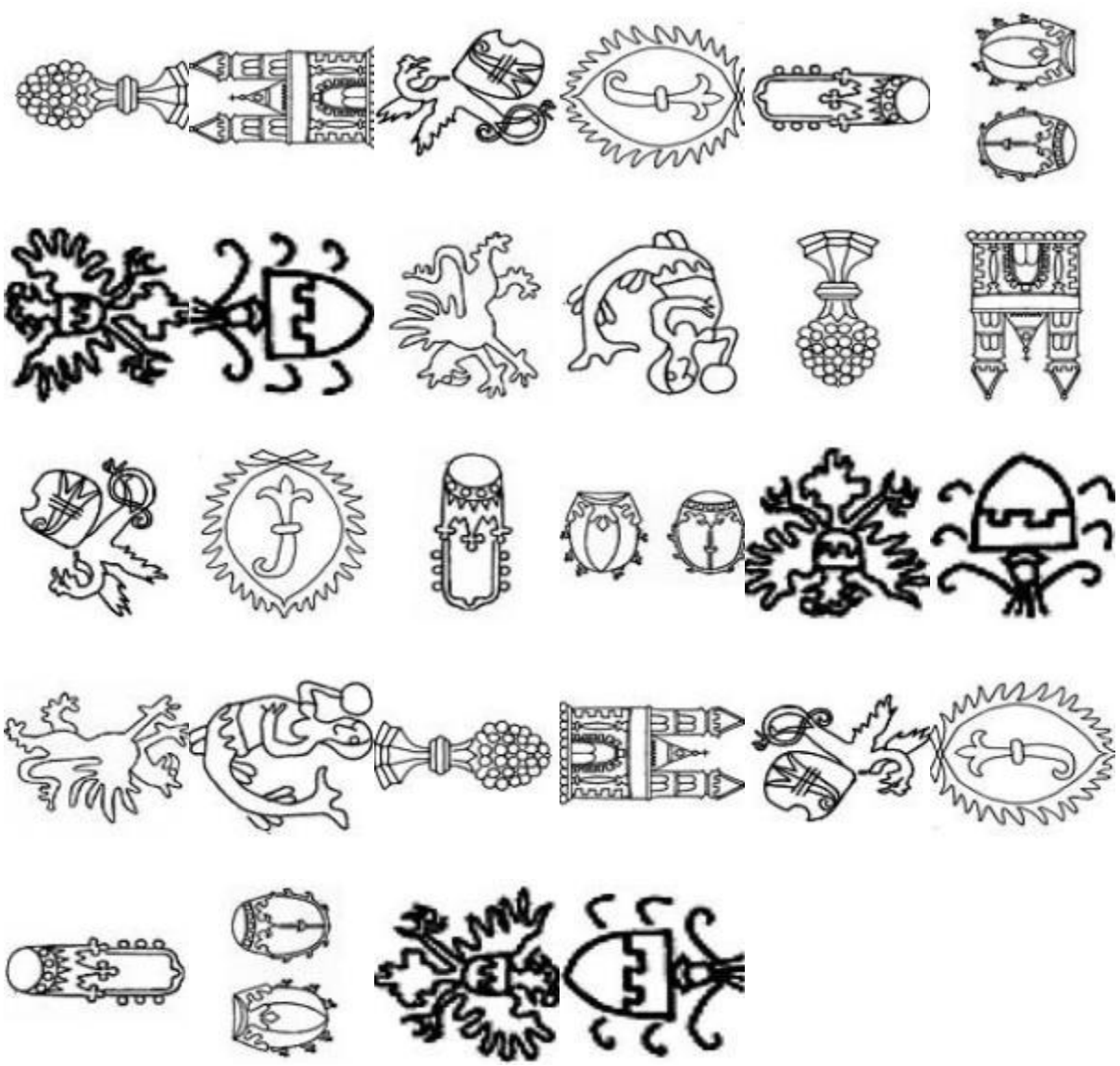










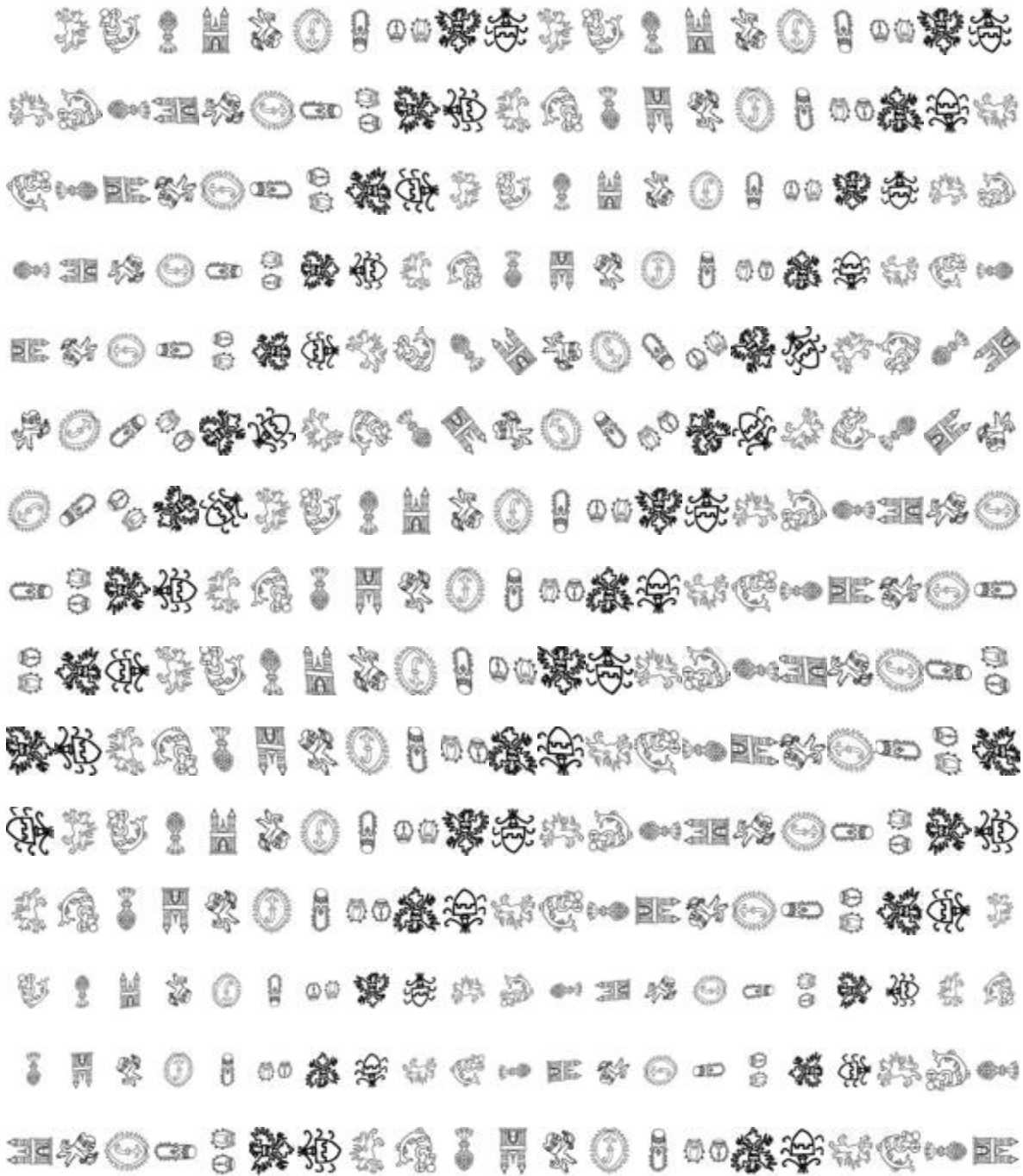




## APPENDIX C

### THE TRANSFORMED AND SCALED IMAGES OF WATERMARKS

The original images of watermarks with size 100 x 100 pixels were transformed and scaled to 28 x 28 pixels images.





## APPENDIX D

### EXAMPLE OF CHAIN CODE

Example of chain code generated for Bernstein watermark #004001001 and its different representations.

The Chain code representation as:

- F8 decimal numeric characters:

```
110077700100766665667770776776766676676666566766766700012222322223222232
2322322232232322221121211212112112226670776776767776767665656565665656665656
6566566665667000001222221222222221221212111111101123222210077011000766665
44566674444432344545555555565656656666666656666701221222212221211101100000
00077776676666566565554455122222344445655667707666654443444443444456666666
6566670070000115666545444566666555443233222234444332232670000012212222222
34444454444444344011011222222334445556667734443443333222222222212111101000
700707777676670122223223222323451223343345554444522322222343222
```

- Binary code with three bits:

```
0010010000001111111100000000100000011111011011011010111011011111111000
111111101111111011110110110111110110111110110110110101110110111110110111
110110111000000000001010010010010011010010010010011010010010010011010010011
010010011010010010011010010011010011010010010010001001010001010001001010001
0100010010100010010100101101101110001111111011111110111111011111111110111
110111110110101110101110101110101110110101110101110101110101110101110110101
110110101110110110110101110110111000000000000000001010010010010010001010010
010010010010010010010001010010001010001010001001001001001001001001001000001001
010011010010010010001000000111111000001001000000000111110110110110101100100
101110110110111100100100100100011010011100100101100101101101101101101101101
1101011101011101101011101101101101101101101101110110110111000001010
01000101001001001000101001001000101000100100100000100100000000000000000000
00011111111111110110111110110110101110110101110101101101101100100101101
00101001001001001001110010010010010111010110111011011111000111110110110110
101100100100011100100100100100011100100100100101110110110110110110110101
110110110111000000111000000000000001001101110110110101100101100100100101110
110110110110101101101100100011010011011010010010010010011100100100100011011
01001001101011011100000000000000001010010001010010010010010010010010011100
10010010010010110010010010010010001110010000001001000001001010010010010010
01001001101110010010010110110111011011011111011100100100011100100011011011
011010010010010010010010010010010010001010001001001001000001000000000111000
```

000111000111111111111111110111110110111000001010010010010011010010011010010010  
011010011100101001010010011011100011011100101101101100100100100101010010011  
010010010010010011100011010010010

- ASCII base 64 representation:

4BA+212/4/3++2+322u32+3ABSSaSTSSaTSaSaTTSSJRRKKJRKW3H+/33/332uuuu1  
12112u122u3AABSSRSSSSRSKJJJBKaSRA/BIA+21ku28kkaclttuuu122221224KRSS  
KSKJIJAAAA//2+212uttktKSSckl1u34+21kjjkkku222123A4ABN21lk1221tkabSSTkkbSa3  
AABSKSSSTkkkkkkjBIJSSsbklt237kjjbbSSSSRRJIA4HH/++3BSSaTSTTIKTcbltkkqT  
SSTjSQA==

- ASCII base 85 representation:

,S0YQ&%:).q#>iXg[4^gmIpEr!D+W\$9dH0%-  
B`("dR]U".&'QfD9)PPp]#[8]'HNN?+&:"[G  
,K+!!#ng"Tu\$5;CJ82,`kW4,XD8?!;gVD]5JsWV^7Y`D7]&@gMOCND7D#/OsqT3+HQ  
R'&-(D  
O[G`q`P1#LbaX,O\*q"no=N\_(IY9=ESIG2\*Q-  
"PpJO8TrimA8/E6:.ZBrXhXoNMFf(WP"B2I8  
E9o\*;Fr4rQJK%<8E=WpP"?p]7\$7fXJ]IfZ[fn%G9dJ2PUtfilA<Nb:8N>-`

## BIBLIOGRAPHY

- [1] D. Hunter, Papermaking: The history and technique of an ancient craft, Dover Publications, New York, USA, 1978.
- [2] R. W. Allison, "An automated World Wide Web search tool for papers and watermarks: The archive of papers and watermarks in Greek manuscripts," 1996, <http://abacus.bates.edu/wmarchive/>
- [3] P. Tschudin, "Papier Geschichte: Sonderausstellung der Basler Papiermuehle," 10 Dez. 1982 bis 31 Mai 1983, Basler Papiermuehle, Museum fur Papier, Schrift und Druck, 1982.
- [4] P. Tschudin, "The evolution of European watermarks," G. C. Castagnari, Editor, Production and use of watermarked papers in Europe, Fabriano, 1994.
- [5] "Watermark Databases," Centre for the Study of the Renaissance, Humanities Building, Rm H448b, University of Warwick, Coventry CV4 7AL, LIMA. <http://www2.warwick.ac.uk/fac/arts/ren/projects/lima/paper/describing/databases/>
- [6] "Survey of known databases of paper structure reproductions," The memory of paper, <http://www.bernstein.oeaw.ac.at/twiki/bin/view/Main/PaperDatabases.html/>
- [7] E. G. Loeber, Paper mould and mouldmaker, Paper Publications Society, Amsterdam, 1982.
- [8] C. M. Briquet, Les filigranes, Dictionnaire historique des marques de papier dBs leur apparition vers 1282 jusqu'en 1600, Tome I B IV, DeuxiBme Cdition, Verlag Von Karl W. Hiersemann, Leipzig 1923.
- [9] C. M. Briquet, "Notice sur le recueil de filigranes ou marques des papiers," presented at the Paris Exposition in 1900. In Briquet's opuscula: The complete works of C.M. Briquet without Les filigranes, pp. 281-288. Paper Publication Society, Hilversum, 1955.
- [10] S. Hawking and L. Mlodinow, The grand design, Random House, New York, 2010.
- [11] M. Kurlansky, Paper - Paging through history, First edition, Norton, 2016.
- [12] Wikipedia, "History of paper," [http://en.wikipedia.org/wiki/History\\_of\\_paper](http://en.wikipedia.org/wiki/History_of_paper)
- [13] N. A. Basbanes, On paper: The everything of its two-thousand-year history, Vintage Books, Random House, New York, 2014.
- [14] A. Monro, The paper trail: An unexpected history of a revolutionary invention, Penguin Random House, New York, 2014.
- [15] D. Woodward, Catalog of watermarks in Italian printed maps ca 1540-1600, The University of Chicago Press, Chicago, 1997, Printed in Italy, 1996.

[16] B. Rudin, Making paper: A look into the history of an ancient craft, Rudins, Box 5058, S-162 05 Vallingby, Sweden.

[17] Wikipedia, "Battle," [https://en.wikipedia.org/wiki/Battle\\_of\\_Talas/](https://en.wikipedia.org/wiki/Battle_of_Talas/)

[18] Bernstein, "The memory of papers project," [http://www.memoryofpaper.eu/BernsteinPortal/appl\\_start.disp/](http://www.memoryofpaper.eu/BernsteinPortal/appl_start.disp/)

[19] J. P. Eakins, A. Jean, E. Brown, J. Riley, and R. Mulholland. "Evaluating a shape retrieval system for watermark images." In Chart Conference Proceedings, A. Bentkowska, T. Cashen, and J. Underland, Editors, Volume four; Digital art history - A subject in transition: Opportunities and problems, British Academy in London, November 28-29, 2001.

[20] D. Francisco de Bofarull Y Sans, Animals in watermarks, Paper Publication Society, Hilversum, 1959.

[21] P.F. Tschudin and E. Riechen, International Association of Papers Historians - IPH, International Standard for the Registration of Watermarks, Provisional Ed., Switzerland, 1992.

[22] "International standard for the registration of papers with or without watermarks," International Association of Paper Historians, IPH, <http://www.paperhistory.org/standard.htm/>

[23] Yale University Library, "Technical art history and conservation research: Watermarks," <http://guides.library.yale.edu/c.php?g=295949&p=1975999/>

[24] Neathery de Safita (aka, Neathery Batsell Fuller), "A brief history of paper," <http://users.stlcc.edu/nfuller/paper/>

[25] N. F. Johnson and S. Jajodia, "Exploring steganography: Seeing the unseen," *IEEE Computer*, 31(2): 26-34, 1998.

[26] F. D. Marmol and N. J. Godenne, Dictionnaire des filigranes classés en groupes alphabétiques et chronologiques, 1900. - XIV, pp. 192, 1987.

[27] C. F. Buhler, Last words on watermarks, Papers of the Bibliographical Society of America, LXVII, 1973.

[28] A. Stevenson, The problem of the Missale special, The Bibliographical Society, London, UK, 1967.

[29] D. Flaut, "Introduction to the auxiliary sciences of history," [www.uuooi.org/english/download.php](http://www.uuooi.org/english/download.php)

[30] S. Spector, Editor, Essays in paper analysis, Folger Shakespeare Library, Washington, USA, 1987.

[31] A. Stevenson, "Shakespearian dated watermarks," *Studies in Bibliography*, 4: 159-164, 1951-2, <http://etext.lib.virginia.edu/bsuva/sb/toc/sib04toc.htm>

[32] R. D. Boyle and H. Hiary, "Watermark location via back-lighting and recto removal," [https://www.researchgate.net/publication/225213127\\_Watermark\\_location\\_via\\_back-lighting\\_and\\_recto\\_removal](https://www.researchgate.net/publication/225213127_Watermark_location_via_back-lighting_and_recto_removal)

[33] A. Stevenson, "Watermarks are twins," *Studies in Bibliography*, 4: 57-91, 1951, <http://etext.lib.virginia.edu/bsuva/sb/toc/sib04toc.htm>

[34] V. Atanasiu, "Ad751 for laid lines density measurement," <http://www.mywebpage.netscape.com/atanasiuvlad/ad751/>

[35] V. Atanasiu. "BlueNile for laid lines suppression and extraction," <http://mywebpage.netscape.com/atanasiuvlad/bluenile/>

[36] V. V. Belov, V. A. Esipova, V. T. Kalaida, and V. M. Klimkin. "Physical and mathematical methods for the visualization and identification of watermarks," *Solanus*, 13: 80-92, 1999.

[37] D. W. Mosser, M. Saffle, and E. W. Sullivan, II. "Puzzles in paper: Concepts in historical watermarks," Papers from the 1996 International Conference on Watermarks at Roanoke, Virginia. Oak Knoll Books and the British Library, New Castle, UK, 2000.

[38] M. Hertzberger, W. A. Churchill. Watermarks in paper in Holland, England, France, etc. in the XVII and XVIII centuries and their interconnection, Amsterdam, 1935.

[39] T. L. Gravell and G. Miller, A catalogue of foreign watermarks found on paper used in America, 1700-1835. Garland Publishing, New York, USA, 1983.

[40] E. Heawood, "Watermarks mainly of the 17th and 18th centuries," Paper Publications Society, Hilversum, 1950.

[41] G. Piccard, Die Wasserzeichenkartei Piccard im Hauptstaatsarchiv Stuttgart: Findbuch, Kohlhammer, Stuttgart, 1961.

[42] A. H. Shorter, Paper mills and paper makers in England, 1495-1800. Paper Publications Society, Hilversum, 1957.

[43] D. C. Waugh, "Soviet watermark studies - Achievements and prospects," *Kritika*, a review of current soviet books in Russian history, Cambridge Press, 6(2): 78-82.

[44] R. W. Allison, "An automated WWW search tool for papers and watermarks: The archive of papers and watermarks in Greek manuscripts." *Puzzles in paper: Concepts in historical watermarks*, pp. 201-210. Oak Knoll Books and the British Library, New Castle, UK, 2000. <http://abacus.bates.edu/Faculty/wmarchive/> and <http://www.watermarkarchive.org/>

- [45] International database of watermarks and paper used for prints and drawings c. 1450-1800, Dutch University Institute for Art History, Florence, <http://www.iuoart.org/wmdb.htm>
- [46] EVTEK, "Paper identification database," <http://conservation.evtek.fi/>
- [47] D. L. Gants, "A digital catalogue of watermarks and type ornaments used by William Stansby in the printing of *The Workes of Beniamin Jonson* (London: 1616)," <http://www.iath.virginia.edu/gants/>
- [48] Hauptstaatsarchiv Stuttgart, "Wasserzeichenkartei Piccard - Piccard watermark collection," <http://www.piccard-online.de/>
- [49] Koninklijke Bibliotheek - National library of the Netherlands, "Watermarks in incunabula printed in the Low Countries (WILC)," <http://watermark.kb.nl>
- [50] D. W. Mosser and E. W. Sullivan, II, "The Thomas L. Gravell watermark archive on the Internet," *Puzzles in paper: Concepts in historical watermarks*, pp. 211-228. Oak Knoll Books and the British Library, New Castle, UK, 2000, <http://www.gravell.org/>
- [51] WIES, "Watermarks in Incunabula printed in Espania," <http://www.ksbm.oeaw.ac.at/wies/>
- [52] WZMA, "Wasserzeichen des Mittelalters," <http://www.oeaw.ac.at/ksbm/wz/wzma2.htm>
- [53] "AHDS: Arts and Humanities Data Service project," <http://www.ahds.ac.uk>
- [54] "MINERVA project," <http://www.minervaeurope.org>
- [55] "PULMAN project," <http://www.pulmanweb.org>
- [56] "EU IST MUSICNETWORK project," <http://www.interactivemusicnetwork.org/>
- [57] "Dylux proofing paper," <http://abacus.bates.edu/Faculty/wmarchive/Dylux.html>
- [58] D. Schoonover, "Techniques of reproducing watermarks: A practical introduction." In *Essays in Paper Analysis*, pp. 154-167. Folger Shakespeare Library, Washington, USA, 1987.
- [59] "T. Gravell," <http://ada.cath.vt.edu:591/dbs/gravell/wizard/wizard.html>
- [60] "DFG (Deutsche Forschungsgemeinschaft, German Research Foundation)," <http://www.dfg.de/>
- [61] DI.MU.SE project Ministero per i Beni e le Attivit`a Culturali and Palatina Library of Parma, <http://www.bibpal.unipr.it/bibliotecaPalatina/index.html>



- [62] “DLF (Digital Library Federation), Digital library standards and practices,” <http://www.diglib.org/standards.htm>
- [63] J. van Aken, “An improvement in grenz radiography of paper to record watermarks, chain and laid lines,” *Studies in Conservation*, 48(2): 103-110, 2003.
- [64] “Beta-radiography method,” <http://aic.stanford.edu/sg/bpg/annual/v01/bp01-02.html>
- [65] “Paper-based watermarks and beta-radiography,” <http://gateway.uvic.ca/lydgate/lydgatems/pages/manuscript/watermarks/watermarks.htm>
- [66] “SHREW: A shape retrieval system for watermarking images project,” <http://www.chart.ac.uk/chart2001/papers/noframes/mulholland.html>
- [67] J. Gancarczyk and A. Olszewska, “Shape based retrieval of ownership and production marks in historical documents,” <https://www.researchgate.net/publication/313222536>  
Shape based retrieval of ownership and production marks in historical documents
- [68] I. Christie-Miller. “Early paper project,” <http://www.earlypaper.com/>
- [121] “Solar Imaging Ltd. APIS (Advanced Paper Imaging System),” <http://www.solar-imaging.com/digital-apis.html>
- [69] “Watermarks of the Middle Ages / Wasserzeichen des Mittelalters, Osterreichische Academie Der Wissenschaften,” <http://www.wzma.at/biblioteken.php>
- [70] “A digital catalogue of watermarks and type ornaments by William Stansby in the printing of the works of Benjamin Jonson (London: 1616),” <http://www2.iath.virginia.edu/gants/Folio.html>
- [71] A. Stevenson, “Paper as bibliographical evidence,” *The Library*, 5th series, 17(3): 197-212, September, 1962.
- [72] F. Del Marmol, “Dictionnaire des filigranes classes en groupes alphabetique et chronologiques,” Namur: J. Godenne, Editor, 1900, - XIV, 1987.
- [73] C. Rauber, T. Pun, and P. Tschudin, “Retrieval of images from a library of watermarks for ancient paper identification.” In *EVA 97, Elektronische Bildverarbeitung und Kunst, Kultur, Historie, Berlin*, Germany, November 12-14, 1997.
- [74] K. J. Riley and J. P. Eakins, “Content-based retrieval of historical watermark images: I-tracings,” Volume 2383, pp. 253-261 of *Lecture Notes in Computer Science*. Springer, 2002.
- [75] J. P. Eakins, A. Jean, E. Brown, Jon Riley and R. Mulholland, “**Evaluating a shape retrieval system for watermark images**,” Institute of Image Data Research, University of North Umbria at Newcastle-upon-Tyne.

[76] W.E.L. Crimson, *Object recognition by computer: The role of geometric constraints*, MIT Press, Cambridge, Massachusetts.

[77] P. Needham, "The study of paper from an archival point of view." In *Historical perspectives in the conservation of works of art on paper*, M. H. Ellis, Editor, The Getty Conservation Institute, Los Angeles, California.

[78] N. E. Ash, "Watermark research: Rembrandt prints and the development of a watermark archive," *The Paper Conservator*, 10: 64-69, 1986.

[79] M. H. M. Van Kessel, "Fast algorithm for the computation of moment invariants," *Pattern Recognition*, 20(6): 639-643, 1987.

[80] K. Jonathan Riley and J. P. Eakins, "Content-based retrieval of historical watermark images: I-tracings." In *Proceedings of the International Conference on Image and Video Retrieval, Lecture Notes in Computer Science*, Volume 2383, pp. 253-261, London, UK, July 18-19, 2002. Springer-Verlag.

[81] K. J. Riley, J. D. Edwards, and J. P. Eakins "Content-based retrieval of historical watermark images: II - electron radiographs." In *Proceedings of the International Conference on Image and Video Retrieval, Lecture Notes in Computer Science*, Volume 2728, pp. 131-140, Illinois, USA, July 24-25, 2003. Springer-Verlag.

[82] G. Salton, *The SMART retrieval system: Experiments in automatic document processing*, Prentice-Hall, 1974.

[83] E. Wenger, V. Karnaukhov, A. Haidinger, N. Merzlyakov, G. van Thienen, E. Oukhanova, and D. Erastov, "A distributed database and processing system for watermarks: An INTAS project." In *Proceedings of EVA'01 Moscow*, pp. 200-206, Centre PIC of Culture Ministry of Russia, STG, Moscow, 2001.

[84] CORDIS, "A distributed database and processing system for watermarks," [http://cordis.europa.eu/project/rcn/65357\\_en.html/](http://cordis.europa.eu/project/rcn/65357_en.html/)

[85] E. Wenger and V. Karnaukhov, "A distributed database and processing system for watermarks." In *Proceedings of EVA'04 Moscow*, The State Tretyakov Gallery, Moscow, CD-ROM, 2004.

[86] E. Wenger, V. Karnaukhov, A. Haidinger, and M. Stieglecker, "A digital image processing and database system for watermarks in medieval manuscripts." In *Proceedings of International Cultural Heritage Informatics Meeting: Cultural Heritage and Technologies in the Third Millennium (ichim01)*, Volume 2, pp. 259-264, Milan, Italy, September 3-7, 2001.

[87] "IPH related databases," <http://www.paperhistory.org/database.htm/>

[88] B. de La Passardi`ere and C. Bustarret, "Profil: An iconographic database for modern watermarked papers," *Computers and the Humanities*, 36(2): 143-169, 2002.

[89] C. Rauber, P. Tschudin, S. Startchik, and T. Pun, "Archival and retrieval of historical watermark images." In *Proceedings of the International Conference on Image Processing (ICIP)*, pp. 773-776, Lausanne, Switzerland, September 16-19, 1996.

[90] C. Rauber, P. Tschudin, and T. Pun, "Retrieval of images from a library of watermarks for ancient paper identification," <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.55.7658&rep=rep1&type=pdf>

[91] C. Rauber, J. O. Ruanaidh, T. Pun, "Secure distribution of watermarked images for a digital library of ancient papers," ACM DL97, Second ACM International Conference on Digital Libraries, Philadelphia, July 23-27, 1997.

[92] P. Tschudin, "Papiergeschichte als Hilfswissenschaft," Sonderdruck aus "Das Papier," 37, Jahrgang, Heft 7, pp. 285-295, 1983.

[93] D. Stewart, R. A. Scharf, J. S. Arney, "Techniques for digital image capture of watermarks," *Journal of Imaging Science and Technology*, 30: 261-267, 1995.

[94] P. Zamperoni. "Wasserzeichenextraktion aus digitalisierten Bildern mit Methoden der digitalen Bildsignalverarbeitung." *Das Papier*, 43(4): 133-143, 1989.

[95] G. Brunner and H. Burkhardt, "Classification and retrieval of ancient watermarks," *Data analysis, machine learning and applications*, pp. 237-244, Springer, 2007.

[96] A. K. Jain and A. Vailaya, "Shape-based retrieval: A case study with trademark image databases," <https://pdfs.semanticscholar.org/d9d2/5ecb12e3f11d5cfd93b4c9d0950ecec22666.pdf>

[97] M. S. Lew, N. Sebe, and J. P. Eakins, Editors, *Image and video retrieval*, In *Proceedings of International Conference, CIVR 2002*, London, UK, July 18-19, 2002, Volume 2383 of Lecture Notes in Computer Science. Springer, 2002.

[98] G. Brunner, "Structure features for content-based image retrieval and classification problems," Dissertation Thesis, Freiburg, Germany, 2007.

[99] T. Han and D. G. Goodenough, "Noise reduction of hyperspectral remotely sensed imagery: A nonlinear dynamical system approach." In *IEEE International Geoscience & Remote Sensing Symposium*, 2008.

[100] X. Briottet, Y. Boucher, A. Dimmeler, A. Malaplate, A. Cini, M. Diani, H. Bekman, P. Schwering, T. Skauli, I. Kasen, I. Renhorn, L. Klasen, M. Gilmore, and D. Oxford, "Military applications of hyperspectral imagery," *Targets and backgrounds XII: Characterization and representation*, 6239:62390B, 2006.

[101] V. Baeten, J. A. F. Pierna, and P. Dardenne, "Hyperspectral imaging techniques: An attractive solution for the analysis of biological and agricultural materials." In *Techniques and applications of hyperspectral image analysis*, pp. 289-307. Wiley, 2007.

- [102] C.-I. Chang, *Hyperspectral imaging: Signal processing algorithm design and analysis*, New York, John Wiley and Sons, 2007.
- [103] Fine Arts Foundation was examined at the University in the Recenart, Research Centre for Art project, <https://www.jyu.fi/en/news/archive/2015/03/tiedote-2015-03-25-20-28-46-577011>
- [104] R. Padoan, T. A. G. Steemers, M. E. Klein, B. J. Aalderink, and G. de Bruin, "Quantitative hyperspectral imaging of historical documents: Technique and application." In *ART Proceedings*, 2008.
- [105] J. Vaarasalo, "Optical properties of paper in papermaking science and technology," *Pulp and Paper Testing*, 17: 162-181, 1999.
- [106] D. V. der Reyden, "Recent scientific research in paper conservation," *The Journal of the American Institute for Conservation*, 31: 117-138, 1992.
- [107] J. De la Rosa and F. J. Bautista, "Optical properties of paper at 337.1 nm." *Revista Mexicana de Fsica*, 51: 110-113, 2005.
- [108] A. Tonazzini, E. Salerno, M. Mochi, and L. Bedini, "Blind source separation techniques for detecting hidden texts and textures in document images." *International Conference on Image Analysis and Recognition*, Lecture Notes in Computer Science, Volume 3212, pp. 241-248, 2004.
- [109] A. Tonazzini, E. Salerno, M. Mochi, and L. Bedini, Document Analysis Systems VI: Lecture Notes in Computer Science, chapter Bleed-through removal from degraded documents using a color decorrelation method, pp. 229-240. Document Analysis Systems VI, 2004.
- [110] A. Belouchrani, K. Abed-Meraim, J. F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Transactions on Signal Processing*, 45(2): 434-444, 1997.
- [111] J. F. Hair, B. Black, B. Babin, R. E. Anderson, and R. L. Tatham, "Multivariate data analysis," Volume 6, Pearson Prentice Hall, 2006.
- [112] M. Attas, "Enhancement of document legibility using spectroscopic imaging," *Archivaria*, 57: 131-144, 2005.
- [113] P. K. Varshney and M. K. Arora, Advanced image processing techniques for remotely sensed hyperspectral data, chapter-Feature extraction from hyperspectral data using ICA, pp. 199-215. Springer, 2004.
- [114] H. F. Grahn and P. Geladi, Techniques and applications of hyperspectral image analysis, Wiley, 2007.

- [115] J. C. Harsanyi and C. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach." *IEEE Transactions on Geoscience and Remote Sensing*, 32(4): 779-785, 1994.
- [116] R. C. Gonzalez and R. E. Woods, Digital image processing, Prentice Hall, Upper Saddle River, N.J., USA, Second edition, 2002.
- [117] A. A. Green, M. Berman, P. Switzer, and M. D. Graig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Transactions on Geoscience and Remote Sensing*, 26(1): 65-77, January, 1988.
- [118] U. A. Diaz, V. M. Reyes, "Determining the dimensionality of hyperspectral imagery," Laboratory for Applied Remote Sensing and Image Processing, University of Puerto Rico, January 2002.
- [119] R. B. Cattell, "The Scree test for number of factors," *Multivariate Behavioral Research*, 1: 245-276, 1966.
- [120] H. F. Kaiser, "The application of electronic computers to factor analysis," *Educational and Psychological Measurement*, May, 1960.
- [121] P. S. Yury, "Continuous measure of significant linear dimensionality of a waveform set," *Computational Statistics and Data Analysis*, June, 2000.
- [122] P. Comon, "Independent component analysis, a new concept?," *Sign. Proc.*, 36: 287-314, 1994.
- [123] T. W. Lee, M. Girolami, A. J. Bell and T. J. Sejnowski, "An unifying information-theoretic framework for Independent Component Analysis," *Computers & Mathematics with Applications*, 31(11): 1-21, Mar. 2000.
- [124] A. Hyvarinen, J. Karhunen and E. Oja, Independent component analysis, New York: Wiley, 2001.
- [125] L. Parra, K. R. Mueller, C. Spence, A. Ziehe and P. Sajda, "Unmixing hyperspectral data," *Advances in Neural Information Processing Systems*, 12: 942-948, 2000.
- [126] J. M. P. Nascimento and J. M. B. Dias, "Does Independent Component Analysis play a role in unmixing hyperspectral data?" *IEEE Transactions on Geoscience Remote Sensing*, 43(1): 175-187, Jan. 2005.
- [127] J. Wang and C. I. Chang, "Independent Component Analysis based dimensionality reduction with applications in hyperspectral image analysis," *IEEE Transactions on Geoscience Remote Sensing*, 44(6): 1586-1600, June 2006.
- [128] M. Lennon, G. Mercier, M. C. Mouchot and L. H. Moy, "Independent component analysis as a tool for the dimensionality reduction and the representation of hyperspectral

images,” *IEEE International Geoscience and Remote Sensing Symposium, IGARSS '01. Proceedings*, 2001.

[129] A. Villa, J. Chanussot, C. Jutten, J. A. Benediktsson and S. Moussaoui, “On the use of ICA for hyperspectral image analysis,” GIPSA-lab, Signal & Image Dept., Grenoble Institute of Technology - INPG, France, Institute de Recherche en Communications et Cybernetique de Nantes, France.

[130] P. Toft, “The radon transform - Theory and implementation,” PhD thesis, Department of Mathematical Modelling, Technical University of Denmark, July, 1996.

[131] M. Sonka, V. Hlavac, and R. Boyle, Image processing, analysis, and machine vision, Thomson, Third edition, 2008.

[132] J. S. Lim, Two-dimensional signal and image processing, Prentice Hall, Englewood Cliffs, N.J., USA, 1990.

[133] N. Otsu, “A threshold selection method from gray-level histograms,” *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1): 62-66, 1979.

[134] J. M. White and G. D. Rohrer, “Image thresholding for optical character recognition and other applications requiring character image extraction,” *IBM Journal of Research and Development*, 27: 400-411, July 1983.

[135] J. Bernsen, “Dynamic thresholding of gray-level images,” *Proceedings of the 8th International Conference on Pattern Recognition*, Paris, pp. 1251-1255, 1986.

[136] R. Cattoni, T. Coianiz, S. Messelodi, and C. M. Modena, “Geometric layout analysis techniques for document image understanding: A review,” Tech. Rep., IRST, Trento, Italy, 1998.

[137] J. Sauvola and M. Pietikainen, “Adaptive document image binarization,” *Pattern Recognition*, 33(2): 225-236, 2000.

[138] P. Viola and M. J. Jones, “Robust real-time face detection,” *International Journal of Computer Vision*, 57(2): 137-154, 2004.

[139] F. Shafait, D. Keysers, and T. M. Breuel, “Performance comparison of six algorithms for page segmentation.” In 7th IAPR Workshop on Document Analysis Systems, pp. 368-379, Nelson, New Zealand, February, 2006.

[140] C.K. Chow, and T. Kaneko, “Automatic detection of the left ventricle from cinean giograms,” *Computers and Biomedical Research*, 5: 388-410, 1972.

[141] W. Niblack, An introduction to digital image processing, Prentice- Hall, Englewood Cliffs, NJ, pp. 115-116, 1986.

- [142] K. V. Mardia, and T. J. Hainsworth, "A spatial thresholding method for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(8): 919-927, 1988.
- [143] T. Taxt, P. J. Flynn, and A. K. Jain, "Segmentation of document images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(12): 1322-1329, 1989.
- [144] S. D. Yanowitz, and A. M. Bruckstein, "A new method for image segmentation," *Computer Vision, Graphics, and Image Processing*, 46(1): 82-95, 1989.
- [145] L. Eikvil, T. Taxt, and K. Moen, "A fast adaptive method for binarization of document images," *Proceedings of ICDAR*, France, pp. 435-443, 1991.
- [146] L. O'Gorman, "Binarization and multi-thresholding of document images using connectivity," *Graphical Model and Image Processing*, 56: 494-506, November, 1994.
- [147] Y. Liu, and S. N. Srihari, "Document image binarization based on texture features," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(5): 540-544, 1997.
- [148] J. Sauvola and M. Pietikainen, "Adaptive document image binarization," *Pattern Recognition*, 33(2): 225-236, 2000.
- [149] T. R. Singh, S. Roy, O. I. Singh, T. Sinam and K. M. Singh, "A new local adaptive thresholding technique in binarization," *IJCSI*, 8(6): 271-277, 2011.
- [150] T. R. Singh, S. Roy, and K. M. Singh, "Local Adaptive Automatic Binarization (LAAB)," *International Journal of Computer Applications*, 40(6): 975-983, February, 2012.
- [151] B. Gatos, K. Ntirogiannis, and I. Pratikakis, "DIBCO 2009: Document image binarization contest," *IJDAR*, 14: 35-44, 2010.
- [152] F. van der Heijden, R. P. W. Duin, D. de Rider, and D. M. J. Tax, classification, parameter estimation and state estimation: An engineering approach using Matlab, Wiley, 2004.
- [153] C. M. Bishop, Pattern recognition and machine learning, Springer, 2013
- [154] C. Tomai, B. Zhang, and V. Govindaraju, "Transcript mapping for historic handwritten document images." In *Proceedings of the Eighth International Workshop on Frontiers in Handwriting Recognition*, pp. 413-418, Niagara, ON, August, 2002.
- [155] E. Saykol, A. Sinop, U. Gadkbay, O. Ulusoy, and A. Cetin, "Content-based retrieval of historical Ottoman documents stored as textual images," *IEEE Transactions on Image Processing*, 13(3): 314-325, March 2004.
- [156] S. A. Shahab, W. G. Al-Khatib, S. A. Mahmoud, "Computer aided indexing of historical manuscripts," *IEEE Proceedings of the International Conference on Computer Graphics, Imaging and Visualization (CGIV'06)*, 2006.

- [157] **D. Ballard and C. Brown**, *Computer vision*, Prentice-Hall, 1982, Chap. 8.
- [158] **E. Davies**, *Machine vision: Theory, algorithms and practicalities*, pp. 149-161 Academic Press, 1990.
- [159] **R. Haralick and L. Shapiro**, *Computer and robot vision*, Volume 1, Addison-Wesley Publishing Company, 1992, Chap. 5.
- [160] **A. Jain**, *Fundamentals of digital image processing*, Prentice-Hall, 1989, Chap. 9.
- [161] K. R. Castleman, *Digital Image processing*, Second edition, Prentice Hall, 1995.
- [162] B. Janhe, *Digital image processing*, Sixth edition, Springer, 2005.
- [163] R. C. Gonzalez and R. E. Woods, *Digital image processing, Third edition*, Prentice-Hall, 2008.
- [164] C. S. Pillai, "A survey of shape descriptors for digital image processing," *IRACST - International Journal of Computer Science and Information Technology & Security*, ISSN: 2249-9555, 3(1), February, 2013.
- [165] A. Amanatiadis, V.G. Kaburlasos, A. Gasteratos, and S.E. Papadakis, "Evaluation of shape descriptors for shape-based image retrieval," *IET Image Process*, 5(5): 493-499, 2011.
- [166] D. Zang and G. Lu, "Review of shape representation and description techniques," *The Journal of Pattern Recognition*, 2003.
- [167] H. Freeman, "On the encoding of arbitrary geometric configuration," *IRE Transactions on Electronic Computers*, EC, 10(2): 260-268, 1961.
- [168] V. J. Hodge, S. O'Keefe, and J. Austin, "A binary neural shape matcher using Johnson counters and chain codes," *Neuro-computing*, 72(4): 693-703, 2009.
- [169] G. Lu, "An approach to image retrieval based on shape," *Journal of Information Science*, 23(2): 119-127, 1997.
- [170] R.O. Duda and P.E. Hart, *Pattern classification and scene analysis*, Wiley, New York, 1973.
- [171] P. Nunes, F. Pereira, F. Marques, "Multi-grid chain coding of binary shapes." In *ICIP'97 Proceedings of the 1997 International Conference on Image Processing*, Volume 3, pp. 114-117, IEEE Computer Society, Washington, DC, USA, 1997.
- [172] E. Bribiesca, "A new chain code," *Pattern Recognition*, 32(2): 235-251, 1999.
- [173] H. Sanchez-Cruz, R. M. Rodriguez-Dagnino, "Compressing bi-level images by means of a 3-bit chain code," *SPIE Optical Engineering*, 44(9): 1-8, 2005.



- [174] H. Sanchez-Cruz, E. Bribiesca, and R. M. Rodrigues-Dagnino, "Efficiency of chain codes to represent binary objects," *Pattern Recognition*, 40: 1660-1674, Elsevier, 2007.
- [175] A. B. M. Salem, A. A. Sewisy, and U. A. Elyan, "A vertex chain code approach for image recognition," *ICGST-GVIP Journal*, 5(3), March 2005.
- [176] H. Freeman, "On the encoding of arbitrary geometric configurations," *IEEE Transactions on Electron Devices*, 10(2): 260-268, 1961.
- [177] X. Dai, S. Khorram, "A feature-based image registration algorithm using improved chain-code representation combined with invariant moments," *IEEE Transactions on Geoscience Remote Sensing*, 37, (5): 2351-2362, 1999.
- [178] H. Sun, J. Yang, M. Ren, "A fast watershed algorithm based on chain code and its application in image segmentation," *Pattern Recognition Letter*, 26(9): 1266-1274, 2005.
- [179] M. Ren and H.A. Karimi, "A chain-code-based map matching algorithm for wheelchair navigation," *Transactions in GIS*, 13(2): 197-214, 2009.
- [180] T. Globacnik and B. Zalik, "An efficient raster font compression for embedded systems," *Pattern Recognition*, 43(12): 4137-4147, 2010.
- [181] N. Ge, G.-Q. Wu, H.-P. Li, Z. Wang, and C.-Y. Bao, "Evaluation of the two-dimensional temperature field and instability of a dual-jet dc arc plasma based on the image chain coding technique," *IEEE Transactions on Plasma Science*, 39(11): 2885-2994, 2011.
- [182] R. K. Gupta and B. Gurumoorthy, "Automatic extraction of free-form surface features (ffsfs)," *Computer Aided Design*, 44(2): 99-112, 2012.
- [183] J. Jain, S. K. Sahoo, S. M. Prasanna, and G. S. Reddy, "Modified chain code histogram feature for handwritten character recognition," *Advances in computer science and information technology, networks and communications*, pp. 611-619, Springer, Berlin Heidelberg, 2012,.
- [184] C. Wang, J. Zhou, P. Kou, Z. Luo, and Y. Zhang, "Identification of shaft orbit for hydraulic generator unit using chain code and probability neural network," *Applied Soft Computing*, 12(1): 423-429, 2012.
- [185] R. Haralick and L. Shapiro, *Computer and robot vision*, Volume 1, Addison-Wesley Publishing Company, 1992, Chap. 5.
- [186] P. Nunes, F. Pereira, and F. Marques, "Multi-grid chain coding of binary shapes," In *ICIP'97 Proceedings of the 1997 International Conference on Image Processing*, Volume 3, pp. 114-117, IEEE Computer Society, Washington, DC, USA, 1997.
- [187] Y. K. Liu and B. Zalik, "An efficient chain code with Huffman coding," *Pattern Recognition*, 38(4): 553-557, 2004.

- [188] Y. K. Liu, W. Wei, P.-J. Wang, B. Zalik, "Compressed vertex chain codes," *Pattern Recognition*, 40(11): 2908-2913, 2007.
- [189] H. Sanchez-Cruz, M. Lopez-Cruces, and H. Puga, "A proposal modification of the 3ot chain code." In *CGIM'08 Proceedings of the Computer Graphics and Imaging*, pp. 6-11, Innsbruck, Austria, 2008.
- [190] H. Sanchez-Cruz, "Proposing a new code by considering pieces of discrete straight lines in contour shapes," *Journal of Visual Communication and Image Representation*, 21(4): 311-324, 2010.
- [191] B. Y.-K. Liu, Z. P.-J. Wang, and D. Podgorelec, "Directional difference chain codes with quasi-lossless compression and run-length encoding," *Signal Processing: Image Communication*, 27(9): 973-984, 2012.
- [192] B. Zalik, N. Lukac, Chain code lossless compression using move-to-front transform and adaptive run-length encoding, *Signal Processing: Image Communication*, 29(1): 96-106, 2014.
- [222] B. Zalik, D. Mungo, and N. Lukac, A universal chain code compression method, *Journal of Visual Communication and Image Recognition*, 29: 8-15, 2015.
- [223] H. Sanchez-Cruz and H.H. Lopez-Valdez, "Equivalence of chain codes," *Journal of Electronic Imaging*, 23(1): 13-31.
- [224] W. Xiaoling and X. Kanglin, "A novel direction chain code based image retrieval." In *Proceedings of the Fourth International Conference on Computer and Information Technology (CIT'04)*, IEEE, 2004.
- [225] S. Junding and W. Xiaosheng, "Chain code distribution-based image retrieval." In *Proceedings of the Fourth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP'06)*, IEEE 2006.
- [226] D. B. Russakoff, C. Tomasi, T. Rohlfing, and C. R. Maurer Jr., "Image similarity using mutual information of regions." In *Proceedings of 8th European Conference on Computer Vision*, 2004, pp. 596-607.
- [227] J. Bing and B. C. Vemuri, "Robust point set registration using Gaussian mixture models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(8): 1633-1645, August, 2011.
- [228] H. Kim, S. Park, S. Lo, J. Monroe, and J. Sohn, "Bidirectional local distance measure for comparing segmentations," *Medical Physics Journal*, 14: 6779-6790, 2012.
- [229] R. Sriraghavendra, K. Karthik, and C. Bhattacharyya, "Fréchet distance based approach for searching online handwritten documents." In *Proceedings of the 9th International Conference on Document Analysis and Recognition (ICDAR '07)*, pp. 461-465, DOI: 10.1109/ICDAR, 2007.121

[230] P. Jaccard, "Étude comparative de la distribution florale dans une portion des Alpes et des Jura," *Bulletin de la Société Vaudoise des Sciences Naturelles*, 37: 547-579, 1901.

[231] M. Levandowsky and D. Winter, "Distance between sets," *Nature*, 234(5): 34-55, 1971. DOI: 10.1038/234034a0

[232] F. Hausdorff, *Grundzüge der Mengenlehre*, Viet. Leipzig, 1914.

[233] S. Nutanong, E. H. Jacox, and H. Samet, "An incremental Hausdorff distance calculation algorithm." In *Proceedings of VLDB Endowment*, 14(8): 506-517, May 2011.

[234] M. Hossain, M. Dewan, K. Ahn, and O. Chae, "A linear time algorithm of computing Hausdorff distance for content-based image analysis," *Circuits, Systems, and Signal Processing*, 14: 389-399, 2012.

[235] W. Rudin, *Principles of mathematical analysis*, Third edition, McGraw-Hill. 1976, ISBN 0-07-085613-3.

[236] A. J. Mansfield and J. L. Wayman, "Best practices in testing and reporting performance of biometric devices," Version 2.01, National Physical Laboratory Report CMSC 14/02, August 2002.

[237] T. A. Louist, "Confidence intervals for a binomial parameter after observing no successes," *The American Statistician*, 35(3): 154, 1981.

[238] B. D. Jovanovic and P. C. Levy, "A look at the rule of three," *The American Statistician*, 51(2): 137-139, 1997.

[239] R. M. Bolle, N. K. Ratha, and S. Pankanti, "Confidence interval measurement in performance analysis of biometric systems using the bootstrap." In *Proceedings of Workshop on Empirical Evaluation Methods in Computer Vision*, Hawaii, December, 2001.

[240] "Confidence interval with zero events," The Monthly Mean newsletter, <http://www.pmean.com/01/zeroevents.html/>

[241] J. Heaton, *Artificial intelligence for humans*, Volume 3: Deep learning and neural networks, Heaton Research, December 2015.

[242] M. D. Zeiler, M. Ranzato, R. Monga, R., Mao, K. Yang, Q. V. Le, and G. E. Hinton, "On rectified linear units for speech processing." In *ICASSP*, pp. 3517-3521, IEEE, 2013.

[243] J. Schmidhuber, "Multi-column deep neural networks for image classification." In *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3642-3649, Washington, DC, USA.

[244] T. Kanungo, R. Haralick, H. Baird, W. Stuezle, and D. Madigan. “A statistical, nonparametric methodology for document degradation model validation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11): 1209-1223, 2000.

[245] T. Kanungo, R. M. Haralick, and I. Phillips, “Global and local document degradation models.” In *Proceedings of the ICDAR*, pp. 730-734, Tsukuba Science City, Japan, October, 1993.

[246] V. Kieu, M. Visani, N. Journet, J. P. Domenger, and R. Mullot, “A character degradation model for grayscale ancient document images.” In *Proceedings of the ICPR*, pp. 685-688, Tsukuba Science City, Japan, November, 2012.

[247] J. Liang, D. De Menthon, and D. S. Doermann. “Geometric rectification of camera-captured document images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(4): 591-605, 2008.

[248] D. G. Lowe, “Distinctive image features from scale-invariant key points,” *International Journal of Computer Vision*, 60(2): 91-110, 2004.

[249] L. Zheng, Y. Yang, and Q. Tian, “SIFT meets CNN: A decade survey of instance retrieval,” *IEEE Journal of Latex Class Files*, 14(8), August 2015.

## VITA

Plamen Doynov received the Diploma Engineer degree in Electronics and Electrical Engineering from Sofia Technical University, Sofia, Bulgaria, in 1985 and the M.E.E. degree in Electrical Engineering from The Catholic University of America in Washington, DC, in 1993. He is currently working towards the Ph.D. degree in the department of Computer Science and Electrical Engineering, UMKC. Plamen Doynov is a Principal Engineer at MRIGlobal, where he is a technical lead for many projects in electrical engineering, software development, and project leader for the design and development of specialize equipment. As a member of multidisciplinary teams working in biology, chemistry and other areas, he has contributed to many turn-key solutions and integrated systems. Plamen Doynov is a Certified Vision Professional and a Certified Biometrics Professional. He is an active member of IEEE (Chair of Kansas City's Biomedical and Computer society chapters, Chair of the Kansas City Chapter). Plamen Doynov is member of Professional Management Institute and is a Certified Project Management Professional. He has more than 25 publications in the areas of Biophysics, Electromagnetics, Neural networks, Brain-machine interactions, Biometrics, and Chemical spectra processing. He has 3 US patents which are granted and one currently in approval process. For contribution to science and technology, Plamen has received twice "Scientist of the year" award at MRIGlobal.