Compact orientation and frequency estimation with applications in biometrics

Biometrics on the orientation express

Anna Mikaelyan
Abstract

Automatic feature extraction still remains a relevant image and signal processing problem even tough both the field and technologies are developing rapidly. Images of low quality, where it is extremely difficult to reliably process image information automatically, are of special interest. To such images we can refer forensic fingerprints, which are left unintentionally on different surfaces and are contaminated by several of the most difficult noise types. For this reason, identification of fingerprints is mainly based on the visual skills of forensic examiners.

We address the problem caused by low quality in fingerprints by connecting different sources of information together, yielding dense frequency and orientation maps in an iterative scheme. This scheme comprises smoothing of the original, but only along, ideally never across, the ridges. Reliable estimation of dense maps allows to introduce a continuous fingerprint ridge counting technique. In fingerprint scenario the collection of irrefutable tiny details, e.g. bifurcation of ridges, called minutiae, is used to tie the pattern of such points and their tangential directions to the finger producing the pattern. This limited feature set, location and direction of minutiae, is used in current AFIS systems, while fingerprint examiners use the extended set of features, including the image information between the points. With reasonably accurate estimations of dense frequency and orientation maps at hand, we have been able to propose a novel compact feature descriptor of arbitrary points. We have used these descriptors to show that the image information between minutiae can be extracted automatically and be valuable for identity establishment of forensic images even if the underlying images are noisy. We collect and compress the image information in the neighborhoods of the fine details, such as minutiae, to vectors, one per minutia, and use the vectors to "color" the minutiae. When matching two patterns (of minutiae) even the color of the minutia must match to conclude that they come from the same identity.

This feature development has been concentrated and tested on forensic fingerprint images. However, we have also studied an extension of its application area to other biometrics, periocular regions of faces. This allowed us to test the persistence of automatically extracted features across different types of images and image qualities, supporting its generalizability.
Acknowledgments

- Uhh, excuse me. If I could have everybody’s attention. I’m Ross Geller.
- Dr. Ross Geller.
- Dad... Please. Anyway, as I was saying, I’m Dr. Ross Geller...

During doctoral study I have been an Early Stage Researchers (ESRs) of the FP7 Initial Training Network BBfor2\(^1\) participating in summer schools and workshops organized by the network partners. Additionally, the network gave me the opportunity to work at guest institutions under various research topics. I would like to thank ESRs as well as seniors for their inputs Ram, Rui, Raymond, Luuk, Antti,... There is no way I can spare everyone’s name, but their input is inevitable. My stays in NL and Finland was most incredible experiences both scientific and culture-wise.

Beyond the words is my dear Intelligent Systems lab, place where I had a lot of laughs and a grate time. I thank the rest of "SAS" for supporting my craziest ideas and for suggesting new ones and for always being there. I thank all the members of HRSS - Deycy for amazing salsa advancing, Adam&Jan- for introducing climbing to our lives. Anita&Jens for being an incredible mentors and being a step ahead to give us easier life. I am grateful to Wagner, Uzma, Asif, Gaurav, Jenifer, Essayas, Suleyman for greater support at and beyond work. I am indebted to Roland for making me go beyond research as to advancing my reporting and modeling skills;) I thank Eva, Anna-Lena and Kristina for making our lives easier and for making it happier. Ksenia, Olessia, Anton and Ania thank you guys, for making my days sunny.

I acknowledge my sincere gratitude for the greater support, my supervisors - Josef, Raymond and Ljudmila who made me thinking and working. Great teachers and researchers they provided the guidance full of patience and enthusiasm, and it was a great honour for me to work with them.

Ivanna& Kolhozniki - I am sad you did not spare time with me during these years. Merci to extraordinary Jane (Sokol and Sherlock naturally), visiting me even with cockroaches and shower on a stairs. I was happy to have Mary here in Halmstad getting with me over the hard days of new Swedish life.

\(^1\)The project Bayesian Biometrics for Forensics (BBfor2) was funded by the EC as a Marie-Curie ITN-project (FP7-PEOPLEITN-2008).
I am thankful to Gert Jan, Joke, Rene, Margot and Bavo and dear Holland for happy days beyond work. I thank my dearest family, who are always there for me, supporting in all possible ways including scientific. Thank you for helping me to get through these years, for patience, belief and your love.

I feel like making an Oscar speech where it feels natural to mention everyone who were there for me and I am lucky to have such people.
## Contents

1 **Introduction** 1
   1.1 Notations and terminology (application only) 1
   1.2 Problem Statement 2
   1.3 List of included papers 4
   1.4 List of supporting papers 5
   1.5 Other relevant publications 6
   1.6 Contributions 7

2 **Ground truth in fingermarks** 11
   2.1 Introduction to Forensic Fingerprinting 11
   2.2 SD27 database of NIST 15
   2.3 Motivation 15
   2.4 Approach Description 17
   2.5 Evaluating performance - CMC and DET/ROC curves 19
   2.6 Discussion 22

3 **Dense maps of frequency and orientation** 25
   3.1 Motivation 25
   3.2 Approach 26
   3.3 Dense vector maps and interactions with experts 37
   3.4 Discussion 41

4 **Extracting features from dense vector fields** 45
   4.1 Motivation 45
   4.2 Approach Description 46
   4.3 Discussion 59

5 **Final Conclusion** 65

Bibliography 69
Chapter 1

Introduction

1.1 Notations and terminology (application only)

Minutia - fine (local) feature of fingerprint image. It is produced by a bifurcation or ending of a ridge line.

Minutia constellation - a set of minutiae data where each minutia data consists in \( (x_j, y_j) \) coordinates and the minutia direction, \( \theta_j \). The more minutiae in the constellation, the more uniquely the constellation can identify a fingerprint.

Core - ridge patterns of parabolic shape in global (coarse) scale

Delta - ridge patterns of hyperbolic shape in global (coarse) scale, with three "asymptotes"

Singularity point (=key point) - generic term used for minutia, Core, and Delta (here). In some studies it is used to refer to Core and Delta (excluding minutia).

Fingermark (=Latent) - trace left unintentionally at crime scene, usually of bad quality and unknown (finger) orientation. The latter refers to the fact that it is not possible to know the exact finger direction from the fingermark, although a qualified guess can be made by the forensic examiner. Fingermarks can be invisible to the eye in which case a chemical processing or dusting is applied to make them visible. Fingerprints of crime scene, are for this reason called latent fingerprints. Orientation maps, minutiae, and minutia directions for such images has been extremely difficult to obtain automatically, if not totally unreliable.

Tenprint - fingerprint, obtained from a controlled acquisition, usually of good quality, and known (finger) orientation. Orientation maps [1] and minutiae can be extracted automatically and reasonably reliably (available as Commercial-Off-The-Shelf Software (COTS) or within AFIS).

AFIS - Automatic fingerprint identification system. System performing database
search to find an enrolled fingerprint image, primarily captured at controlled image acquisition locations, e.g. embassies, border control posts, national identity card offices, custody/jail. Such fingerprints are traditionally taken by inking and rolling the finger from nail edge to nail edge on a card (fingerprint card) or they are captured electronically by a live scan device. They are called tenprints as all ten fingers are typically captured. Evidently, tenprints have a good image quality. Input for AFIS is however not an image in forensics, but minutia constellations, possibly including cores and deltas locations and directions.

Identification - Examiner extracts fingermark minutiae and sends it to central AFIS system which automatically selects a short-list (e.g. top 20 matches) from hundreds of thousands fingerprints. Then fingerprint expert checks if the query image is really in the short-list and if so then files a report on the technical basis for her/his visually verified features (local, global) additionally checked by another fingerprint expert.

1.2 Problem Statement

Information quality in signals remains one of the key problems of automatic processing of data, despite the increased complexity of sensing systems capturing the data. There are plenty of situations when more complex sensing systems have not (yet) been able to deliver more informative knowledge. In this work we will touch upon some of the challenges of biometric image processing, especially in connection to forensics.

Fingerprints have been suggested and used in identification and verification since 19th century. While in daily biometric applications error of recognition rushes to zero, in forensic framework there is still a room for significant improvement. To be identified, a partial area of an unknown finger with an unknown (finger) rotation having noise of different sources requires great deal of manual work.

In the current state-of-the-art analysis, the examiner extracts all fine characteristics, the number of which may vary from 1 up to 200. Such characteristics are primarily minutiae, cores and deltas, collectively called singularity points here. While extracting singularity points the examiner visually analyzes the fingerprint pattern both locally and globally, though the only information she can insert into the identification system is locations and directions of singularity points. The difficulty arises when the number of such features is low, typically less than 10, because geometric locations (two integers each) and directions (an integer each) are not sufficient to link a fingermark to a finger beyond doubt. Minimal number of minutia points required for identification varies from country to country and is between 7-17 (Russia - Italy) points [2].
1.2. PROBLEM STATEMENT

Other countries do not have a numerical standard, essentially relying on the human expertise.

As a result, our goal will be to provide a tool to translate expertise of the examiner to a machine readable code or, in other words, encode image information into a feature that can be understood by a human and used by an automatic matcher. In practice, we are interested in defining and computing additional properties from neighbourhoods of singularity points, such that their information content (currently location and direction) can be increased. What patterns will these features represent? Will the patterns have an implicit reference point that can be linked to the location of the singularity point? Will the additional information reduce the need for a high number of singularity points, in other words, how much identity can the additional information buy when singularity points are scarce in the fingermark?

The main goal of our work in iris biometrics was to verify the generalizability of the feature that was developed for fingerprints. Unlike point-based features used in fingerprints which have poor image quality, iris images have high quality and are rich in texture. Accordingly, state-of-the-art features used to identify iris, measure texture properties. Texture features seek to give unique identities to (local) regions because textures are repetitive by definition. By contrast, our features measure properties that aim to give unique identities to points. Additionally, iris biometrics has larger annotated databases of images that are publicly available, than forensic fingerprint biometrics. Bigger data provides better statistics to evaluate performance and generalizability, which motivates our choice of iris biometrics as additional application.

Both theoretical framework and practical application of theory are important pieces of knowledge. Therefore, we have set the goal to benefit both, on one side building the context which helps to understand why and how the technique is working and when it is failing. On the other hand, we wanted to make sure that a theory is not built up for an ideal vacuum situation, but is tightly connected to a problem, to which a tangible contribution is needed. As such, we have contributed to the following fundamental problems of theory and practice.

i Evaluation is essential to any scientific contribution. Nonetheless, publicly available annotated datasets providing ground truths are scarce in forensic fingerprints. The annotations that exist do not always permit to evidence progress or provide knowledge in sufficient depths. If available, more detailed ground truths would improve the depth and breadth of the progress.

ii Reliable way of estimating dense frequency (scale) and orientation images is one of the fundamental image processing problems, e.g. in object recognition, enhancement or measurement of objects.
iii Image processing is lacking a generalizable compact image descriptor of objects, holding strong identity and quality information with reference to a unique location. The available features are either not well studied or extract texture measures which are unique with reference to regions (rather than points).

iv Forensic images possess extreme amount of noise hampering automatic image processing. At the same time examiners must do much of the monotonic work, increasing the risk of human error, e.g. in feature extraction, ridge counting, matching. Reliable automatic image analysis tools, if available, would improve the quality of identification by reducing the human time in the process, while increasing the amount of the technical evidence, (extracted by machines but verified by humans).

These problems are coupled on different levels of image analysis. As such, the theoretical problem iii) requires a stable image information as an input therefore contributing to the solution of ii) was a necessary phase. Similarly, to solve iii) it was important to obtain ground truth i).

1.3 List of included papers

The following papers are appended. Detailed description of the papers along with author contributions will follow below and in subsequent chapters, Ch. 2-4.


The first author is the main contributor of the paper – active participation to discussions, formation of the outcomes, and the documentation.


The contributions of the second author is significant, near half if such matters would be numerically quantifiable, and are detailed as follows. Sections 2-6: active participation to the discussion, formation of the outcomes, and the documentation, although the contributions of the second author did not exceed those of the first author. Sections 1, 7-9: active participation to discussions, formation of the outcomes, and documentation. The contributions of the first author did not exceed those of the second author.
1.4 List of supporting papers

The following papers are reviewed and published. They are, except a subset of [D1], comprised in the included papers, therefore omitted. Paper [D1] has an element that is not comprised by [A], [B], [C], namely fusion, which is peripheral to the thesis, though supportive. The letters indicate which appended paper the papers are included in.


Novel ridge counting method based on dense frequency and orientation information was developed in order to validate the potential of dense maps in forensic fingerprint identification. Ridge counting is a part of the experimental setup of paper [B].


The paper presents a novel feature extraction algorithm applied to forensic fingerprint images. Feature vectors are extracted from a dense orientation maps. Experimental results are tested for core points of the images in which they exist and constitute a subset of the experimental setup of paper [C].


SAFE features which were introduced earlier, [C1], are applied here to dense frequency maps in addition to orientation maps. Experimental results are extended from describing cores to minutia points lying in neighbourhoods of the latter. The results are included into the journal submission [C]. The paper has been selected for excellent paper award in ISCIT 2014.
The novel feature descriptors suggested for low quality forensic fingerprints are now applied to periocular images. The evaluation of performance of suggested object-based feature and its complementarity to a texture-based Gabor feature is embodied into submission [C].

Paper introduces further work comprising comparison of SAFE features to other methods in application to periocular and iris images, and fusion.

1.5 Other relevant publications

Below is the list of papers that are not reviewed or their reviews are not finished, relevance of which to the enclosed papers is denoted with letters. Paper [E1] provides further evidence to generalization of dense frequency and orientation maps, in addition to iris biometrics. It is not included in supporting papers nor appended to the thesis because, as of writing, its review is not completed.


1.6 Contributions

As a main contribution, we identify a development of novel feature extraction method for highly noisy images. Not only the method is developed in the mathematical framework allowing for a theoretical ground of the features but also it is applied and tested on real images across different applications.

Development of a new method of image description requires existence of ground truth. In the search of ground truth we analyzed the only publicly available database of forensic fingerprints, NIST SD27 database [3], comprised of 258 pairs of images, one acquired from a crime scene and the other taken in a supervised environment. Current fingerprint matching techniques consider two images same if all the singularity points and their directions match. In order to support the research beyond state-of-the-art, we have established ground truths of correspondences at singularity point level, and studied baseline performance of publicly available matchers, as detailed in Paper A.

Before constructing features it is important to reduce the amount of noise in forensic fingerprint images. We suggested a novel procedure for obtaining automatic dense orientation and frequency maps for fingerprint images with high level of noise. The estimation is done iteratively via frequency map, orientation map, and enhancement of the original by adaptive Gabor filters, as precised in Paper B.

Having both ground truth information about the singularity point correspondence and enhanced orientation images we have stepped into the final state of suggesting a novel feature descriptor. Filters were designed in such a way that they attempt to extract the image information in the ring shaped neighbourhoods of a singularity point in a compressed form. This information encodes changes of orientation around the point, projected onto space of harmonic functions. The latter define the location of a singularity point uniquely, because all harmonic basis functions, except one, have one and the same singularity, defined implicitly, the origin (where the minutia, core or delta is located). Furthermore, we provide evidence to the generalizability of the description power of the suggested features by stepping out of the forensic fingerprint application. We have applied them to recognition of periocular and iris images. The results and the details of the experimental evidence on their usefulness are reported in Paper C.

The contributions are listed in further detail next with applicable category – theory (t) and practice (p).

(t1) Novel way of constructing automatic dense orientation and frequency maps, (See paper [B], Section VIII for description of the iterative scheme where one map is estimated from the other, Sections II-IV for the detailed construction of maps at each step of the iterations and Section 3.2 of the thesis for summary).
General model-based compact feature descriptor of image information in neighbourhood of a key point (Paper [C], Section III A-C and Section 4.2 of the thesis). Our descriptor vector, representing the orientation/frequency information around the point compactly, contains built-in quality estimate while it is rotation invariant (Paper [C], Section III D-E where rotation compensation and quality estimation are explained and Section 4.2 of the thesis).

Ground truth for forensic fingerprint database. We have established correspondence at minutia level for NIST SD27 database (Paper [A], Section IV). This is significant for understanding, development, and in depth evaluation of methods for fingermark images, e.g. for establishing baseline performance at minutia level (Paper [C], Section IV-B and Section 4.2 of the thesis).

Continuous ridge counting technique for forensic fingerprints, and interpolating missing information (Paper [B], Sections VII, IX-A and Section 3.2 of the thesis).

Point-wise matching of fingerprint images (Paper [C], Section IV-B and Section 4.2 of the thesis).

Experimental support for the suggested object-based descriptor being complementary to texture-based Gabor features (Paper [C], Section IV-C and Section 4.2 of the thesis).

Detailed contributions (contribution letters a, b, c express the connection to papers [A], [B] and [C] correspondingly)

(a1) Study of verification abilities of two publicly available minutia constellation matchers on SD27 database in terms of EERs and CMC curves (Section III of the paper and Section 2.4 of the thesis).

(a2) Experimental support of poor performance of constellation based matchers in forensic scenario justifying further research to increase their information content (currently location and direction). We detail the contribution in Section IV of the paper and in Section 2.4 of the thesis.

(b1) Automatic estimation of filter parameters for improved orientation map in application to fingerprints (fingermarks as well as tenprints). We detail the contribution in Sections IV-A, IX of the paper and Section 3.2 of the thesis.

(b2) New way of estimating the local spatial frequency without knowing orientation nor the enhanced image first. The latter are obtained in an iterative process after frequency estimation (Section VIII of the paper and Section 3.2 of the thesis).
1.6. CONTRIBUTIONS

(b3) Frequency estimation is shown to be orientation estimation but in logarithmic scale space where the total least square optimal estimates are given (again) by the structure tensor (Section IV-B of the paper and Section 3.2 of the thesis).

(b4) Ridge counting method based on frequency and orientation maps without explicit ridge localization (Sections VII, IX-A of the paper and Section 3.2 of the thesis).

(b5) Evidence of the dense orientation and frequency maps in vicinities of singular points being a useful descriptor of a minutia, cores and deltas (Section IX-B of the paper and Section 3.2 of the thesis).

(c1) Novel model-based feature extractor with rotation invariance and built-in quality estimates measuring properties sensitive to locations of singularity points (object/particle properties as opposed to texture/wave properties). We detail the contribution in Section III of the paper and Section 4.2 of the thesis.

(c2) Design and mathematical construction of dense filter functions, with its two spatial support properties (location of the rings and their relative widths/overlap) being geometrically steerable (Sections III A-B, IV-A and Appendix of the paper and Section 4.2 of the thesis).

(c3) Experimental support for object-based SAFE feature being complementary to texture-based Gabor features (Section IV-C of the paper and Section 4.2 of the thesis).

(c4) Reporting both identification and verification performance addressing different needs of comparisons, including repeatability (Sections III, IV of Paper [A], Section IV-B of Paper [C] of the paper and Section 2.5 of the thesis).

(c5) Recognition experiments in isolation from other features showing identification power of features alone (Sections IX-B of Paper [B], Section IV-B of Paper [C] of the paper and Section 4.2 of the thesis).

In addition, we contributed to answer some relevant questions (q) of forensic and image analysis sciences

(q1) What makes the location of a point unique? We think that a possible answer is how much its neighborhood has particle-like properties, i.e. whether or not the point is a singularity (in the mathematical sense) of its neighborhood.

(q2) A related question is "What makes a point anonymous?" We think that how much the point is part of a wave determines how difficult it is to locate (a wave, like a texture, is not localizable) its constituent points.
(q3) How many points is enough for identification? For fingerprint identification the answer is contributed to by extending the feature space of a point, from currently being the direction (i.e. the minutia direction) to include symmetry properties of the image around the point. Even though the exact number is not identified, it is tightly connected with the first two questions and the uniqueness of the point and its neighbourhood.
Chapter 2

Ground truth in fingermarks

2.1 Introduction to Forensic Fingerprinting

"It often seems to me that’s all detective work is, wiping out your false starts and beginning again." Yes, it is very true, that. And it is just what some people will not do. They conceive a certain theory, and everything has to fit into that theory. If one little fact will not fit it, they throw it aside. But it is always the facts that will not fit in that are significant." 
(Agatha Christie, Death on the Nile)

Fingerprint skin surface forms ridges which are noticed to have unique patterns (W. Hershel, 1858) and used in identification or identity verification of humans (and sometimes animals). Since late nineteenth century (1882, A. Bertillon) fingerprints are used as evidence in criminal investigations. In addition to unique pattern, fingerprints barely change over time and can differ even in identical twins, unlike DNA. First Identification Division (for criminal investigation) was established in 1924 by the Department of Justice of the United States, but courts started accepting fingerprint as scientific evidence even earlier in late 1800’s in India and Argentina (1911 in US and 1912 in Sweden) [4].

The uniqueness of fingerprint patterns was not scientifically or numerically proven, but at this time courts took an important decision concerning fingerprint identification: "It has occurred to us that instead of the state being called upon to offer proof that no two fingerprints are alike, it may now be considered in order for those taking the opposing view to assume the burden of proving their position" [5]. The decision was based upon years of forensic experience as well as medical scientific research. The latter shows that formation of friction pattern is influenced by genetics as well as the overall environment during the development of the unborn baby. According to the law of biological
variation "Nature never repeats". Once the friction pattern is formed, it stays permanent and unique [6].

Crime scene investigators often search for fingerprint impressions, that can possibly be left by the perpetrators. They use a variety of techniques, including powders and chemicals to discover and extract this, often invisible, impression to be apparent to an eye. Investigators explore the areas where humans often leave prints. Sometimes invisible trace found at a crime scene is named latent, while the developed image is called fingermark in Europe and, again, latent in US, Fig. 2.1 Right.

Technician’s work on latents results in images, here and further referred as fingermarks, which are passed to the examiners, who analyze the prints for further identification. Due to the predominantly low quality of these images, examiners manually explore the images to extract fine, characteristic details which originate from disruption of a fingerprint ridge pattern, called minutiae, Fig. 2.5 Left. The minutia points are also referred to as Galton details, according to the British anthropologist who identified them in the first book on fingerprints. In the book he divided all possible impressions into four types: arches, loops, whorls and composites, which were later used by a police inspector E. Henry for classifying fingerprint images. At the same time all impression types have so called fixed points which are core and delta, comprising the aforementioned fingerprint patterns, Fig. 2.2.

To match a fingermark, fingerprint examiner marks positions and directions of minutia points, cores and deltas and submits the query to an identification software, performing a one-to-many search. The database which examiner
2.1. INTRODUCTION TO FORENSIC FINGERPRINTING

searches is usually securely located and consists of fingerprints acquired for criminal or noncriminal justice purposes. The impressions are made either with ink or with a scanner, allowing for automatic "lights-out" processing and minutiae extraction. The fingerprints are commonly referred to as tenprints since all ten prints are usually acquired from the person. The minutia set, marked by the examiner is compared to thousands of automatically pre-marked minutiae sets of images stored in the database.

First, an automatic matching software using a database, known as Automated Fingerprint Identification System (AFIS), was developed in 1980 by FBI Identification Division [8], Fig. 2.3. On input it accepts a set of minutia points marked by the examiner and as output it generates (usually) 20 best matches found in the database. The examiner visually inspects the suggested matches via side by side comparison of the fingerprints and their respective singularity points. If all singularity points are judged to agree, then the examiner declares identification, otherwise she declares elimination, [9]. To reduce possible human error, two examiners perform the same identification process independently and confirm the final (yes/no) decision with each other. This step finalizes the consequence of actions ACE/V (Analysis, Comparison, Evaluation, and Verification) of identification. The comparison can be done using magnifying glass and can include direct comparison of two fingerprints, excluding database search if a suspect already exists. To perform identification, it is important that the measurements applied to fingerprints are recognizable and/or verifiable by humans.

Even though no two fingerprints were found alike, there has been a case where subparts of them, found at crime scenes, were. The famous landmark is Daubert’s case (1993), after which the court started if not doubting, but considering admissibility of fingerprints. Scientific and medical research of fin-

Figure 2.2: Left: Core (circle) and delta (square) of a fingerprint. Core and Delta are coarse (global) scale singularities. Often, distinctive features are nearby cores and deltas. If such singularities are available in fingerprint, the probability of identification is higher. Middle: Fingerprint of an arch type. Right: Fingerprint of a loop type, [7]
Figure 2.3: State-of-the-art matching using constellation of minutiae i.e. location and direction information per minutia. Analysis step of ACE-V algorithm. Minutiae of a fingermark (Top) are extracted manually by a forensic examiner, while minutiae of the tenprint are extracted automatically by a machine (Bottom). The number of minutia between a fingermark and tenprint differs significantly. In addition, fingermark images possess higher amount of noise.

Gerprints formation shows the uniqueness and permanence of friction skin, [10]. Unfortunately, scientific foundation is lacking for forensic fingerprints, where only partial areas are available, [11]. We still struggle to answer both which areas to use and how many minutia points uniquely define a fingerprint. Quantification of the weight of partial fingermarks in casework is the continuing tendency in forensic fingerprinting [12], [13]. Authors have evaluated the influence of rare and not rare minutiae on matching, though minutia rareness is estimated manually by the examiner.

Recent years forensic identification tends to move away from testifying in terms of absolute certainty (match/no match) towards using a language containing statistical probabilities and odds when making identification. The probability is based on the frequency of appearance of fingerprint pattern in the population. This probability reflects the statistical odds of fingermark, which is often just a part of a full print, belonging to a third party other than defendant and is called likelihood ratio. Uniqueness of fingerprint in whole does not prompt the uniqueness of its subparts. However, in combination they can provide for a strong uniqueness for a fingermark.
2.2 SD27 database of NIST

The database was developed by NIST and contributed by FBI. It contains two sets of images, one from crime scenes (fingermarks) and one acquired in controlled environments (tenprints). Each image has two sets of minutiae (four in total for a tenprint/fingermark pair), encoded by trained forensic examiners. These sets are called Ideal and Matched set. The Matched sets of a mated tenprint and fingermark are then two sets of minutiae, each containing some minutiae from the respective images. The number of minutiae are the same in Matched sets of a tenprint–fingermark pair because each minutia has a corresponding minutia in the other image. There are no minutia with non-matching minutia in the other image. Ideal sets of a tenprint–fingermark pair have generally different numbers of minutiae each. This is because there are matching as well as non-matching minutia, Fig. 2.4. It is called "Ideal" because the set that an examiner extracts from a fingermark and the automatically extracted minutia set from its matching tenprint have different sizes due to there exist matching as well as non-matching minutiae. To be able to declare identification the examiner needs to explain the minutia that are not present in the common part of both images.

Annotated minutia sets correspond thus to a procedure performed by the examiner, as at the first step he/she extracts all visible (guessable) minutiae in the image, some of which later might be discarded. First set is extracted without expert seeing the other image (Ideal set) and the second with minutiae that exist on both of the matching images after having seen both (Matched set). Final visual verification of common minutiae is done by the examiner. In addition to minutiae sets, examiners have also labeled cores and deltas of fingerprints if they exist (also Ideal and Matched).

In accordance to the quality of fingermark images, all pairs were split into three subsets - Good, Bad and Ugly. The subsets are roughly equally populated. It is not mentioned in the database description how the quality of the image is defined but the number of minutiae comprised in a fingermark (those that examiner can extract with confidence) is low for ugly, average for bad and high for ugly sets.

The database consists of 258 pairs of fingerprints, a fingermark and a tenprint. Images of fingerprints exist in a higher resolution too, but, we have studied 500 dpi (50 micron pitch) resolution since it is the common format in which tenprints has been stored in law-enforcement agencies. It is also an FBI standard format for AFIS, while the higher resolution images are more common for third level features, namely sweat pores.

2.3 Motivation

It is a capital mistake to theorize before one has data.
CHAPTER 2. GROUND TRUTH IN FINGERMARKS

Figure 2.4: Comparison step of ACE-V process where both minutiae sets of a tenprint/fingerprint pair originating from the same finger (mates) are given. Due to non-uniform force application and nonlinear distortion the locations and directions of minutiae may differ between the two impressions slightly. This, and the different qualities of the two impressions causes undetected minutiae or falsely detected minutia, as illustrated by the ideal set, by the human expert (fingerprint) and the machine (tenprint).

(Sir Arthur Conan Doyle., A Study in Scarlet)

It is common in biometrics to use a large constellation of minutia points to identify a person while in the forensic scenario the number of such points may be too low and not enough for a definite verdict. We want to increase the discrimination power of state-of-the-art features, which is a position and direction per minutia point, with an orientation change information derived from the (image) neighbourhood around the minutia, see Fig. 2.5. As first step we need to find how much useful information the new features add to the existing features. This can help clarifying whether many minutia points are necessary to be able to match a fingerprint with a tenprint. We need the ground truth to tell if a certain image processing results in a true improvement of an identity recognition. Though it seemed to have annotation and thereby ground truth, the labeling of minutiae in SD27 is regretfully at finger (=person) level, whereas we wished them be at minutia level. To this end, we had to establish the minutia pair labels for the minutiae of the database. Additionally, we had to establish the recognition performance of the currently used features (minutia location+direction) to appreciate what novel features would offer.
2.4 Approach Description

Forensic fingerprint data is difficult to obtain due to ID protection and considerable (human) resources which are required to create annotation. For non trained human eye it is a problematic task to annotate such a database.

We performed the following steps to obtain the ground truth at minutia level. First, we used publicly available software (with source code) to establish minutia-wise correspondence. The methods perform matching using only the state-of-the-art features, (minutia location + direction). By reediting the software we were able to obtain what such a method suggests in terms of minutia correspondence (which were visually verified later), that is the we have studied its performance beyond verification scores (Paper [A]).

To our knowledge this is the first reporting on matching minutiae sets annotated by fingerprint experts at minutia level using a published matching method (with publicly available implementation). These matchers (known as k-plet [14] and bozorth3 [15]) have verified fingermark–tenprint pairs using the similarity between their minutiae constellations. We will call those matched subsets Supposedly Matched. We have measured performance in terms of Receiver Operator Characteristics, ROC, and Cumulated Match Characteristics, CMC curves. Both curves are detailed in the paper and commented further next. The rank-1 (correct) recognition rate summarizes a CMC curve, and this was 78 % for k-plet and 66% for bozorth3. Likewise Equal Error Rate, EER, summarizes a ROC (or its cousin DET curve). Performance of k-plet is 36% in EER on Ideal and 6% on Matched sets. The error is much lower for matched minutia sets since the fact that there are nearly no non-matched minutiae ren-

Figure 2.5: Left: The two minutia types are exemplified - bifurcation and ending of ridges. Both types have a direction which is tangential to the relevant ridge direction. These directions are visualized with a vector starting at the minutia location. Middle&Right: The suggested features are extracted from ring shaped areas, e.g. the one between the two circles, around a singularity point. The way defining directions differ from country to country and even between the organizations within the same country, e.g. the directions of FBI and NIST in USA differ by $\pi$. Above and in the rest of the paper we use the NIST standard.
Figure 2.6: **Left:** Tenprint minutiae labeling. **Middle:** Labeling of the corresponding minutiae in the mating fingerprint by (automatic) matchers using minutia constellations. **Right:** Corrected minutiae labeling in the same fingerprint. Note for example the top right minutia 14 (Middle, Blue) is now labeled as 10 in correspondence with tenprint.

Second, we have established the ground truth visually and manually, by correcting the errors of the matcher itself. We wrote displaying and editing software to facilitate this error-correction by means of a human intervention. The correspondence at minutia level when minutia are displayed on images is visually obvious to verify for a human eye, but for a machine this is still a major challenge.

After the second phase, the *Permuted Matched* set was obtained for 258 pairs of SD27 and each minutia was given a unique label (number) within a matching pair. Altogether, there were 11 minutiae that were erroneously annotated (due to human expert error, e.g. maximally incompatible minutiae orientations - 180°). Therefore, we marked the 8 fingerprint–tenprint pairs and did not use them in our experimental evaluations. Final set contains 5449 visually verified and reordered minutiae such that the identities of the minutiae, encoded in their storage order (and label), correspond.

It was not touched in Paper [A], but we have also sanitized the correspondence of cores and deltas between Ideals and Matched sets of fingerprint and tenprint. There were 36 (7) identities for which transfer of cores (deltas) occurred between match and ideal sets. Some fingerprints had Ideal sets that were not larger in population than their corresponding Matched sets, which contradicts the definition of the Ideal set. We have consequently marked and imported the corresponding minutiae from Match sets in a hope that examiners verified Match sets more carefully in comparison to Ideal sets. After sanitation...
2.5 Evaluating performance - CMC and DET/ROC curves

Images analysis of forensic fingerprints lies in the intersection of forensics, biometrics and image processing. Accordingly evaluation of results is also special in some aspects but also common in others. Due to the procedure of fingerprint analysis used by forensic examiners, the latter awaits from the matching software to output approximately 20 images to compare them to the input fingerprint. Therefore, if the matched pair exists in the database it is important and labeling of cores and deltas, Matched set of the database has 57 (114) identities that had no cores (deltas) and 201 (144) that had cores (deltas).

Figure 2.7: Types of erroneous minutia annotation (left - fingermark, right - tenprint). Top: Missing minutia (42) in the tenprint while existing in the fingermark. Bottom: Differing orientations in minutiae that correspond (20).
that it pops up among the first 20 suggested by the matching software. At this stage, for a forensic expert it is more natural to see the curve which visualizes the probability of the matched tenprint to be in the first N, Fig. 2.8 Right. Such curves are called Cumulative Match Characteristic (CMC) curves.

Cumulative Match Characteristic (CMC) curve is a measure of performance of identification (one-to-many). CMC is obtained from observations of lists of sorted tenprint candidates upon fingermark queries, the ones having the highest matching score with the enrolled image having the highest ranks

$$\text{CMC}(k) = \frac{1}{n} (\#k_l \leq k), \; l = 1 : n$$

(2.1)

where $\#k_l$ is the number of the mating tenprints observed at rank $k_l$ and $n$ is the size of the database, [16]. It represents the probability of correct identification up to different ranks.

The queries are assumed to be on closed set, i.e. the tenprint of a queried finger is assumed to exist in the database, which in itself is problematic because it puts the burden of proving that the assumption is false on the human examiner. Such curves depend heavily on the size of the database, hampering comparison of different matching algorithms unless the used databases have the same size and similar diversities. The difficulty of comparison is illustrated by Fig. 2.9 Left where it is evident that the skills of the shooters are not possible to assess solely based on the shots being closer to the center. Even the background conditions, e.g. the shooters distance to the targets must be taken into account. Similarly, a CMC curve displays the performance that does not provide a fair comparison basis between the algorithms if the database size, from which the tenprints must be recalled, is not the same between the algorithms. This is confirmed further by [17], see Sec. 5.1 and 19.

Receiver Operating Characteristic (ROC) or Detection error trade-off (DET) curves in turn measure the performance of verification (1:1). Curves plot False Reject (FR) vs. False Accept (FA) rates together visualizing the performance
2.5. EVALUATING PERFORMANCE - CMC AND DET/ROC CURVES

Figure 2.9: Shooters at the different distances to the target, hampering comparison of results using targets only. This is analogous to CMC curves where the results cannot be used to compare the skills of matchers if the background conditions (database size and composition) are not the same.

as the two error types are traded off against each other. False reject is a case of missed detection of the tenprint mate upon a fingerprint query while false accept represents pulling a false tenprint (=innocent) as the mate. ROC curve can be used to designate an acceptable operation point for a matcher, since while moving on the curve one error type always increases as the other decreases.

DET curve is an refurbished version of ROC, essentially carrying the same information. The main difference is that a non-linear mapping, similar to logarithmic, is applied to both FA and FR axes [18]. The authors of the latter paper have suggested using the observed impostor scores and client scores distributions (rather than Gaussians) to construct the mapping. DET curves thereby intend to depict how much the FA and FR score distributions resemble normal distributions, in addition to illustrating the performance trade-off between the two fundamental error types of a given algorithm. The more a DET curve is similar to a line, the more likely the FA and FR scores come from normal distributions. As in ROC curves, it is possible to extract the EER from a DET curve (the intersection point of the DET with the 45 degrees line).

The Equal Error Rate (EER) point measures the separation power of the matching algorithm, namely how well impostors differ from client images using a standard operating point (where FA and FR are equally bad to have for an operator). EER is established as an intersection of FRR and FAR. All curves (FA versus scores, FR versus scores, ROC, DET, and CMC) can be extracted from verification scores if both impostor and client tests are undertaken in well designed verification experiments. The client tests represent then fingerprints being matched against their respective mating tenprints and impostor tests
CHAPTER 2. GROUND TRUTH IN FINGERMARKS

Figure 2.10: FA and FR curves are obtained as integrals of client and imposter score distributions (in practice their observations in a series of verification experiments) by moving the threshold on the score axis, as indicated. The intersection point between the resulting FA and FR curves is the Equal Error Rate. This is when the two integrals above become equal at a certain threshold and represents the goodness of a matching algorithm in making correct decisions when pulling the mating tenprint upon a fingermark query.

represent verifications when every fingermark being tested against every other (non-mating) tenprint.

When a fingermark expert evaluates the quality of her own decision, e.g. in case of a match, she is interested in odds of having made the wrong decision. At that point, the examiner wishes to know the distributions of whatever measures she has used in the general population. Thereby, even a human examiner is interested in the ROC curves in the process of decision making, but at a later stage, because odds are readable from figures like Fig. 2.10, for (every) threshold the examiner uses in her decision.

2.6 Discussion

The other goal was to establish a baseline performance of state-of-the-art open-source matchers (as opposed to commercial matchers which do not disclose method details), at finger as well as minutia level, on the only publicly available annotated database.

Even though Paper [A] presents mainly experimental findings, the ground truth finding would be not possible to accomplish without scientific methodology, i.e. careful redesign of existing software not meant to be used for matching at minutia-level. This has allowed us to obtain results on how publicly available software would fare, if the whole ground truth was established completely automatically, highlighting the challenges that remain for fingermark-tenprint matching.

As first step of the ground truth establishment, we determined the performance of state of the art published minutiae matchers on fingermark-tenprint
problem, which is about 40% EER as was mentioned above. After correcting human errors in the original annotation, the error rate went down significantly. Recent finding with the updated version of algorithm bozorth3 and k-plet show EER of 21% on Ideal (Sanitized) set.

Our sanitization guarantees that a match set is a proper subset of an ideal set, which was not the case in the original annotation, presumably due to that experts were given incoherent instructions on how to obtain the matched sets. Several possibilities exist as to how the Matched sets were obtained: i) an expert sees a genuine tenprint/fingerprint pair and the corresponding ideal sets but allowed only to do deletion from the Ideal sets to obtain the Matched sets, ii) ditto but the expert is also allowed to add (novel!) minutiae, and iii) a Matched set is obtained by seeing only the image pair, without seeing the Ideal sets. Only production strictly following i) will yield a Matched set which is a proper subset of an Ideal set. Even if only the Ideal set is used in algorithm development it is useful to know which minutia correspondences it found were actual ones and which actual ones were missed. Additionally, the researchers can choose to turn on or off minutiae from the common sub-set or non-common subset to evaluate and improve their algorithms. This cannot be done without knowing the proper subset that is common to the two Ideal sets.

The size of the database does not allow for extensive experimental setups used in testing commercial systems, where larger amounts of sequestered data (not available publicly) are used. On the other side, it is not a commercial or fully automatic system that is the target of our study, but to produce tools that will help a human expert to make better visual investigations. The likelihood of finding a similar fingerprint by a coincidence is quite high for 258 images. This explains why human visual system reigns supreme in fingerprint.
Chapter 3
Dense maps of frequency and orientation

3.1 Motivation

To uniquely identifying a key point we wish to rely on its neighbouring image information. To this end the image should to be as noise-free as possible in the neighborhood. Preliminary observations on the SD27 fingermark database showed that there is a potential to give an "identity" to a key point by using its image neighborhood only, i.e. without using the minutia constellation of the fingermark. For single "good" minutia of a fingermark we were able to pull out the corresponding minutia (in the matching tenprint) to be in the first 20 (out of 5449 minutia of all tenprints). Unfortunately the number of such minutia with high quality surrounding is low. This motivated that a new enhancement technique focusing on the quality of the extracted features equipping key points with identity or discriminatory information.

To cope with "noise" in forensic image analysis we take the view that such images have pixels with "vector" values instead of the traditional gray values, bringing the possibility to represent a multitude of visual properties of pixel neighborhoods, and the quality at which they can be observed therein, Fig. 3.1. The analogy is color where each pixel of a color image can be assumed to be a vector, although the color vectors are sensed (e.g. by a color camera) whereas the components of our vector are computed from the image neighborhood of a pixel. Our vectors are thus dense feature vectors where a part of each vector represents the quality of the measured visual property, [19]. Forensic experts pay a great attention to the quality of the regions in which they estimate/locate various visual landmarks. Accordingly we think that there is a need to quantitate the observability or the "quality" of a measurement even
technically, especially since the use of the term "quality" in relation to evidence presented in courts have been on the rise during recent years.

In fact, using color to display dense local orientation and frequency images (also called maps) have been used in image processing before. These maps are used in higher level operations yielding motion estimation, noise reduction, detection of corner points, texture classification, and texture segmentation, permitting a long array of appreciated applications, e.g. face recognition in cameras of mobile phones, iris recognition in biometric identification, medical imaging, object tracking, optical character recognition, etc. Orientation and frequency in fingerprints are relatively easy to define since these correspond to the direction and the spatial frequency, respectively, of the ridge flow around a point. Initial motivation for targeting to improve orientation and frequency maps instead of the enhancement of the original fingerprint, is to suppress non-relevant observable orientations from relevant orientations, which is often a problem in fingermark images.

Below, when Fourier transforming sinusoids under Riemann integrals, their integrability would become an issue, since they are not absolutely (Riemann) integrable, and their Fourier transforms exist only in the sense of (Dirac) distributions. Accordingly, to handle distributions, the integrations below should formally be interpreted as Lebesgue integrals.

3.2 Approach

Noisy images are not a specific forensic problem, it is a challenge of the image analysis in general. We address it by connecting different sources of information together yielding dense frequency and orientation maps in an iterative scheme comprising smoothing of the original, but only along (not across) the ridges of a fingerprint. Other higher level measurements directly depend on the computed orientation and frequency maps. Ridge distance computations, for example, are influenced by frequency as well as orientation of ridge lines.

Usage of structure tensor (ST) for constructing orientation maps is not novel [1], but automatic tuning of the filters used to create good quality STs for every pixel, and thereby a high quality of orientation image, is. We do this iteratively. To expose how this is done, we need to study the computation of structure tensor below, see also Fig. 3.2.

i) At first we convolve image \( f \) with the (complex) filter \( h \): \( f_{\text{grad}} = h \ast f \)

\[
h(x, y, \sigma_{in}^2) = (D_x + iD_y)g = (D_x + iD_y)\frac{1}{2\pi\sigma_{in}^2}e^{-\frac{x^2+y^2}{2\sigma_{in}^2}}
\]

where \( \sigma_{in}^2 \) fixates the inner-scale (the frequency-contents) resulting in an image with complex valued pixels, \( f_{\text{grad}} \).
3.2. APPROACH

Figure 3.1: The goal of our enhancement is to obtain a dense orientation (and a dense frequency map, not shown) of a good quality, and not an "eye pleasing" enhancement of a fingermark to replace the original. This is because higher level measurements depend on such maps more directly than they do on the original. In the color image the intensity is the quality of the local orientation (bright=good) whereas the hue is the local orientation. Note that hue is an angle, in color spaces where Red, Green, and Blue are assumed to be orthogonal (HSV color space is used). We use enhanced images to obtain such maps, but we expect that if our enhanced image at an intermediate state is bad, the estimated quality (of the orientation) will reflect this by being low in such regions (black/dark in the color).

ii) Apply point-wise (complex) squaring to obtain infinitesimal linear symmetry

\[ f_{ils}(x, y) = f_{cgrad}^2 \]

iii) Convolve \( f_{ils} \) and \( |f_{ils}| \) with a Gaussian filter \( g(x, y, \sigma_{out}^2) \) defining the neighbourhood (the outer-scale)

The maximum magnitude in Step i) is obtained when the maximum and minimum of a derivative filter coincide with a valley and ridge of fingerprint respectively. The derivative filter \( h \) is the complex version of the gradient of a Gaussian with variance \( \sigma_{in} \). It can be shown that the distance between the maximum and the minimum of such a filter (real or imaginary part) is \( 2\sigma_{in} \), in pixels. The three steps, comprising two LSI systems injected with (non-linear, pixel-wise) squaring in between, result in a frequency band to which the ST computations are most sensitive. Spectrally, the result of Step i) is equivalent to that the Fourier transforms of \( f \) and \( h \) are multiplied, \( F \cdot H \). However, because a Gaussian rapidly decreases with increased frequency, even if it is multiplied with a polynomial of any order (spatial derivations), there is an effective band associated with every \( h \). Thus, amplitudes of \( F \) outside of the effective band determined by \( H \) are suppressed/ignored.
CHAPTER 3. DENSE MAPS OF FREQUENCY AND ORIENTATION

Figure 3.2: Top: The first step of Structure Tensor computations is a Linear Shift-Invariant System, LSI. The LSI System is represented as a black box, transforming an input sinusoidal signal to an output sinusoid keeping the frequency $\omega_0$ intact but amplifying the amplitude and adding a (constant) phase. In the first step the standard deviation of the used filter, $\sigma_{in}$, determines which frequencies $\omega_0$ get amplified most. A fingerprint pattern is represented by a sinusoidal wave. Step ii) of ST computations comprise pixel-wise squaring/multiplication of the results of Step i). Step iii) is again an LSI system which is a smoothing of the results of Step ii) with a filter parameter $\sigma_{out}$.

Because filtering with a symmetry derivative of a Gaussian, which is what eq. (3.1) is, [20, 21], is linear (Linear Shift-Invariant System), the input sinusoidal signal, by which a local fingerprint ridge/valley image can be modelled, results in a sinusoid too. Accordingly, if the input signal has the frequency $\omega_0$ which is in concordance with the peak frequency of the filter, then the output signal magnitude reaches a peak value.

The tensor, so called because it is a matrix that measures an observable physical property (orientation) of the signal, unbiased by the coordinate system of the image. That is, even if the camera taking the image is rotated ST will calculate the orientation of the image of the corresponding (physical) neighborhood with exactly same accuracy. Parameters $(\sigma_{in}, \sigma_{out})$ are important parameters of the structure tensor defining its scale, with $\sigma_{in}$ fixating the inner-scale, i.e. the frequency contents and $\sigma_{out}$ fixating the outer scale, i.e. integration component defining the locality. By increasing the window size $\sigma_{out}$ of a filter one can reduce the amount of noise but at the cost of diminishing spatial resolution.

For tenprints (which typically have good quality) one can use predefined values for $\sigma_{in}$, and $\sigma_{out}$, which can be determined heuristically based on experiments relying on tenprints. However, for fingermarks (and other low quality images), "one size does not fit all". We need to adapt these parameters to the
Figure 3.3: Squaring introduces sinusoid terms in places (higher part of the spectrum) they did not exist but also amplifies the lowest frequencies. **Left Top:** Fourier transform of a signal. **Left Bottom:** Fourier transform of the squared signal. **Right:** Applying Gaussian to the squared signal suppresses the introduced high frequencies and sharpens the lowest frequencies. If the valuable information is in the low frequencies, it can be extracted reliably despite the introduction of high frequencies, often considered as noise.

image at hand, and even to sub-regions of images if the application demands it. However, informative regions of fingermarks left on crime-scenes are usually small, meaning that there is a chance $\sigma_{in}$, and $\sigma_{out}$ are constant throughout the image, albeit these are unknown a priori. In that case the same parameters can be used to estimate all STs in the image and thereby the dense orientation map (including the quality). In Paper [B], we have established a method to estimate these parameters from a fingermark at hand, automatically. That is, each time we have a fingermark image, (or its enhanced version of it), we (re)estimate these filter parameters to obtain orientation fields from STs that are best adapted to the current image. It is shown in the paper that the outer scale parameter, $\sigma_{out}$, completely depends on the smallest allowable absolute frequency ($\approx (\pi/2)\omega_{\text{min}}$) in the image, eq. (20) of Paper [B]. Thereby, it depends on the inner scale parameter $\sigma_{in}$, because it is also shown that the latter depends on which effective (absolute) frequency range is present in the image, Multiple Orientation Theorem of Paper B. A fixed $\sigma_{in}$ may be therefore good enough for a fingermark because the frequency range cannot be very large in small areas representing such images but as said it must be estimated from the fingermark at hand.

We introduce the orientation estimation procedure as the iterative scheme comprised of (see Fig. 3.4): Dense frequency estimation $\Rightarrow$ Dense orientation estimation $\Rightarrow$ Image enhancement $\Rightarrow$ Dense frequency estimation $\Rightarrow$ ...
procedure is repeated until convergence in (global) mean frequency is achieved. An advantage of the suggested frequency estimation is its estimation of the absolute frequency without a need to know the underlying orientation. This allows to construct a dense frequency map directly from the input image only, which in turn allows to estimate the filter parameters estimating the orientation map more accurately. Subsequently both maps are used to enhance the original image by adaptive smoothing which relies on a correct estimation of the frequency and the orientation.

We retain the orientation map and the frequency map of the last cycle and use them for further processing, e.g. for feature extraction and for ridge counting. An enhanced image is also obtained at every cycle, including the last, but this is currently a byproduct which is not used directly in our feature extraction albeit it is used indirectly, because the dense maps depend on all of the intermediary enhancements and dense maps in combination. The enhanced image cannot be proposed to forensic experts because, they prefer to use the original image to not be biased.
3.2. APPROACH

The frequency estimation is the initial computation because this allows an accurate average frequency estimation for a fingermark. Even if the allowable frequency range is known, the optimal frequency which is the most commonly occurring frequency is not known, which is what determines the choice of the filters used in the orientation map estimations. Thus, frequency estimation must be possible to estimate without knowing the orientation.

In paper [B], we go further than estimating the average frequency. We estimate a dense map of it, from which we estimate the average. In turn, dense frequency estimation is constructed by fitting orientation to a logarithmic scale space of the structure tensor trace (paper [B]). That is, this requires at least computation of a discrete scale space for $\log(I_{11})$ which is achieved by computing $\log(I_{11})$ at 3 different inner scales $\sigma^2$ chosen around a $\sigma^2$ determined by the global mean frequency (via $\sigma^2 = 1/\omega^2$). This scale-space is non-linear w.r.t. the image, $f$, because already $ST, <\nabla f(\nabla f)^T>$, is nonlinear w.r.t. to $f$, and thereby the logarithm of its trace is nonlinear in $f$.

Dense absolute frequency maps are also used as inputs to enhance fingerprints, which is done by smoothing the original input along fingerprint ridges. Prior studies have shown that such anisotropic smoothing, adaptive to ridge orientation and frequencies, is superior to traditional approaches of denoising techniques relying on isotropic filtering, e.g. [22]. The smoothing could have been implemented by adaptive filtering using local frequency tuned (Gabor) filters [22] in (local) blocks of the spatial domain, or in the Fourier domain, as in [23], [24], using the energy spectrum itself as a filter. This would entail heavily down-sampling of the orientation map compared to the original size of the fingermark. Our experiments have confirmed that more informative orientation and frequency maps are obtained if no down-sampling is made. This is explained by the fact that the frequency changes vividly in the neighborhoods of minutiae, which are still the most important clues of identities. Down-sampling of rapid changes cause then significant information loss. Instead we have computed a novel Gabor filter for every pixel of the fingermark and projected the original to this. The ensemble of projections yields smoothing along the ridges.

Image enhancement by projections on Gabor filters, [22], or its kin quadrature mirror filters [25], do not constitute a novelty in themselves. Nor is image enhancement using structure tensor is new, [26]. Nonetheless, the theoretical findings on the urgency of estimating absolute frequency sufficiently well and how to couple the estimation to filter-size (as detailed in the paper), which helps to reduce heuristics in image enhancement by Gabor projections, are novel. To be precise, the calculated Gabor filters are novel, Paper [B] Appendix, because previous studies did not hint on how the Gabor filter size (not the frequency of the complex-sinusoidal part) along and across the current ridges are related to the estimated frequencies.
Master orientation and master frequency maps are obtained as those in the final cycle, when mean global frequency has converged. Typically it requires 2-4 cycles for convergence to occur for SD27 fingermarks.

The paper offers a mathematical unification of orientation estimation and absolute frequency estimation, namely that both problems reduce to orientation estimation problem. We believe that this is a novelty relevant to systematization of image analysis. A practical consequence is that with few scale-space samples it is possible to obtain useful absolute frequency estimation along with a certainty. The certainty has then a physical interpretation, total least square fitting of an orientation to scale-space. In automatic scale estimation or interest point estimation, e.g [27], [28], the original sampling grid of the scale parameter limits the accuracy of the location of scale-space extrema. Interpolation must then be employed to obtain higher accuracy. Confidence in automatically estimated scale, in the frame of scale normalized derivatives and physical interpretation of the confidence is not a well studied matter in published studies.

In Paper [B], we have introduced another use of absolute frequency maps in fingerprint processing, to count ridges continuously along an arbitrary path, without detecting ridges explicitly. The computation needs dense orientation and frequency maps, nothing else. It is demonstrated that already 3 scales in fingerprint ridge counting yields absolute frequency estimations good enough to target application, ridge counting.

Structure tensor trace, as driving force to generate a Gaussian based non-linear, isotropic scale space is a construction that has not been studied well previously, although the tensor has been applied to frequency decompositions, e.g. [29], [30]. If linear scale space is to be used to measure absolute frequency then at least second order derivatives, which can be made isotropic (e.g. Laplacian), are needed. This study shows that a non-linear construction of scale space saves a factor of $\sqrt{2}$ of multiplications per metric dimension, i.e. factor 2 for 2D images. The absolute frequency is estimated isotropically by magnitude responses of first order derivatives. The latter are anisotropic but the magnitude responses of a vector/complex filter composed of all derivatives of Gaussians are isotropic. Thereby, the suggested non-linear logarithmic scale space offers an original way to quantify image properties related to scale using more compact filters than those of linear scale space theory, used in automatic scale estimation.

Another issue that we have studied is the isotropy in connection with the outer scale, and the size of the outer scale itself. The isotropy and the size of the neighborhood in which the local maximization is made, [27], is not well studied in automatic scale estimation by using the linear scale space. It is reasonable to couple it to a-priori available information, as it is done here and elsewhere, but then there is no-guarantee that the max operation will be able to extract the local amplitude/energy in the signal isotropically, even if Laplacian is isotropic, and the window max operation is circular (i.e. as circular as a
3.2. APPROACH

digital grid allows). The subsequent processing, including absolute frequency or scale estimation, risks then to be biased by the orientation of the (local) image. This is shown by using double sinusoids in all possible orientations, and changing the local period linearly, paper [B].

The main mechanism of frequency estimation in Paper [B], as well as previous studies, e.g. [27], includes pulling out the frequency information by observing the response of the local image to a series of linear filters whose frequency characteristics are known, i.e. probing by known "instruments" and observing the result. Our probing happens in the first step of the ST calculation by filtering the original with a complex gradient, eq. (16) Paper [B]. At this point the sought quantity is to be found in the amplitude of the output sinusoid, because it amplifies it in a way which is characteristic to the input frequency, but also in the (output) sinusoid term itself. By way of example, this probing or interaction happens via a series of Laplace filters in [31] and the frequency information is then "hidden" in the same way as in our method, i.e. in the amplitude and in the sinusoid term. The sinusoid term is the undesired term in both methods because the value of a sinusoid changes with image location whereas frequency, which is a wave property is desired to remain constant throughout all image locations where the wave is present.

In our method we use pixel-wise squaring in the second step of ST calculation, which does two useful things. First, it transfers the useful term, the amplitude after interaction with the probe filter, to the lowest frequency possible (zero frequency or DC component). Second, the unwanted ripple term which originally had the frequency $\omega$ is now "transfered" to a sinusoid with double frequency, $2\omega_0$, as illustrated by Fig. 3.3. Thereby, the subsequent smoothing which happens in the third step of the ST calculation suppresses the ripple term more efficiently than not squaring (and smoothing). Needless to say, the DC term containing the useful measurement comes out "unharmed" and constant throughout all image locations where the input wave is present. Evidently squaring demands that the original image is sufficiently densely sampled such that squaring does not cause aliasing. If this presupposition is suspected not to hold, the original can simply be doubled in size appropriately (zeros in every second row and column and appropriate low-pass filtering) as a preventive measure.

By contrast the extraction of the useful term and suppressing the ripple term in [27] happens by a maximum operation in a "sufficiently" large neighborhood (the size of which is not specified). While this works reasonably well for signals without noise, Paper [B] shows that the maximum procedure is more vulnerable to noise than squaring and smoothing.

In paper [B], we have evaluated four additional frequency estimation techniques from theoretical and practical view points. These are on one side Hong [22], and Chikkerur [32], which were suggested in biometric identity recognition applications and are well studied w.r.t. (recognition) performance, but have succinct theory or/and poor generalization ability (to other applications). On
The other side, we have also studied Knutsson [25], and Lindeberg [27] which were suggested as general image analysis tools, having well studied theory, though not having performance statistics in biometrics, Fig. 3.5.

In the paper we have tested the improvement of enhancement by using a minutia extraction (mindtct), and a matching (bozorth3) techniques, both implemented and documented by [15]. Open source and third party implementations are important for repeatability and reduction of risks of biased implementations. In the experiments, we have compared the frequency map of [22] and [32], and the one we have suggested in a processing chain where everything else remained identical. The chain has comprised matching manually extracted minutia (by experts) from fingermarks with automatically extracted minutia from enhanced tenprints.

Additionally, we have suggested a new method of ridge counting between two minutia points. The method performs ridge counting continuously without a need for explicit localization of ridges. It relies on availability of dense orientation and frequency maps. The ridge-counting method is also evaluated to verify that the suggested frequency and orientation maps can be used for this novel purpose, and that our dense maps offer advantages over a state of the art generalist maps, [27], when used for fingermarks.

The ridge count on an arbitrary path between two points (including minutiae but not limited to) uses a model where the image neighborhoods on the path are represented by a planar wave, \( \cos(\omega_0 r) \), with \( \omega_0 \) being the wave vector (frequency of the image ridge flow calculated beforehand) and \( r \) being a point on the path.

In the special case, when the ridge orientation \( \angle \omega_0 \) is constant and the path is a line (Fig. 3.7 Left Red), we have \( N_L = \frac{L}{T} = L \frac{\hat{\omega}}{2\pi} \), where the period and the frequency along the line (red in the figure) are related to each other as \( \hat{T} = \frac{2\pi}{\omega} \), the distance between the points is \( L \). Then the frequency in the direction of the (red) line is \( \hat{\omega} = \omega \cos \alpha \) where \( \omega \) is the frequency of the ridge.
3.2. APPROACH

Figure 3.6: Correct and erroneous ridge count estimation examples. Most of erroneous estimations arise from either recipient ridges or from minutia misplacement.

wave along the gradient (with $\omega = \|\omega_0\|$) (yellow in the figure), and $\alpha$ is the angle between the line and the ridge gradient (yellow in the figure, orthogonal to the ridge direction).

In practice the orientation on the inter-minutia path is not constant. However, the equation of number of ridges is still valid but on (infinitesimally) small segments of the inter-minutia line, meaning that it can be calculated by a line-integral

$$N_L = \frac{1}{2\pi} \int L \hat{\omega}(x,y)dL = \frac{L}{2\pi} \int q(\tau)\omega(\tau)|\cos(\varphi(\tau)) - \varphi_L(\tau)|d\tau$$ (3.2)

The frequency $\omega$ can be obtained from the frequency map, whereas the quality $q(\tau)$ can be a combination of the quality obtained from frequency and orientation maps, e.g. their products. Here $\varphi(\tau) - \varphi_L(\tau) = \alpha$ is, as above, the angle between the direction of the tangential direction of the path and the gradient direction of the ridge-wave (available via the orientation map). The theory is developed in the continuous domain assuming a finite number of sinusoids as input, which is the case for fingermarks (which are even smaller in size than tenprints).

Automatic enhancement of the orientation and frequency maps resulted in 78% ridge count estimation agreement (within 1 ridge error on +70000 minutia pairs) between fingermark and tenprint using all possible minutia pairs and 83% ridge count estimation agreement (within 1 ridge error on +6000 minutia pairs) between fingermark and tenprint when pairs with edges passing by other minutiae are removed from statistics.
In computing ridge count feature, some combinations of minutiae are prone to error, for example, when the minutiae directions (or the opposite direction) are close to the line connecting the two minutiae Fig. 3.7 Right. We implemented an automatic rejection scheme for such lines, using orientation maps. We have studied this error source, (called second type of constellation errors) in conjunction with the ridge counting method.

Being critical when selecting paths on which ridge counts is made is already built into work flow of forensic experts, when such counts are needed. When using ridge counts as evidence, it is more intuitive and practicable for experts to go from one minutia to the next by visiting the nearest minutia, rather than attempting to count ridges on every possible minutia edge, thereby evading edges being too close to other minutia. Likewise, human experts avoid choosing paths running parallel to ridges. Additionally, It is also normal for forensic experts to avoid minutia edges passing regions with no ridges, simply because they cannot count them. Accordingly, an automatic system for counting minutia can suggest edges to a human expert, and human expert can reject some of the suggested edges interactively, bringing in her valuable experience, and improving the quality of her assessment by machine counts.

Finally, the ridge counting section of the paper links frequency and orientation maps including their certainties to continuous ridge-counting, comprising interpolation when image quality is poor, or even lacking partially. In the latter case a propagation of average frequency and orientation estimations,
3.3 Dense vector maps and interactions with experts

Drawing and colour are not separate at all; in so far as you paint, you draw. 
The more the colour harmonizes, the more exact the drawing becomes. 
Paul Cezanne

Forensic expert may not be directly helped with automatically obtained 
orientation map, if she is not habituated to how to interpret them. We think 
that habituation problem can be conveniently solved by displaying an orient-
tation map side by side with a circular color reference as exemplified in Fig. 
3.8. Even such a simple presentation (displaying a reference chart) can quickly 
eliminate the problem of associating a color with its reference orientation rea-
sonably precisely and an unhabituated expert can start using her knowledge 
to correct the errors of orientation, if need be.

Evidently, even more advanced, specially developed Graphical User Inter-
faces, GUIs, for Human Machine Interactions, HMI, can be envisaged, software 
programming resources allowing. For example, the expert can click on a point
and a zoomed rectangle is shown on the same image (or in an image of its own beside) where the color and the original image are blended gradually. Alternatively the color and a synthetic ridge pattern (estimated by means of the orientation map, the frequency map, and the Gabor phase) can be blended dynamically until the original is produced, etc. Such GUIs, although very useful, were not implemented given that the present study had to focus on the fundamentals of image processing.

In all our (appended and supporting) papers we have extensively used color images to represent angle information. Accordingly dense orientation maps were presented as color images as above when we were analyzing/verifying orientations of fingermarks. The same representation was also used in computers: one vector per pixel where the magnitude of the vector always representing the quality of the estimation. [19]. However, in state-of-the-art orientation maps are heavily down sampled and human analysis of them is made by bars as shown in Fig. 3.8. As the color rectangles indicate, it is difficult to display, verify and correct the rapid changes of orientation, e.g. around cores and deltas, by a human since a drawn (black-white) bar occupies many more pixels than a (color) pixel.

Color representation is especially natural when using Linear Symmetry Tensor (LST), an analogue of ordinary structure tensor but using complex numbers. How this complex representation is coupled to displayed/printed orientation maps is discussed next. We begin with the complex number

$$LST = (I_{20}, I_{11}) = \left( \int ((D_x + iD_y)f)^2, \int |(D_x + iD_y)f|^2 \right)$$

where $I_{20}$ represents the directional part and $I_{11}$ is a scalar measuring the nondirectional part of the local image. Linear symmetry direction present in $I_{20}$, defines the direction of a local image along which it has a unique direction for all of its isocurves, e.g. parallel ridges, a single line, or an edge. A unique isocurve direction in this context is the direction along which the local image varies the least upon a small translation.

ST (eq. (2) of Paper [B]), as a symmetric matrix is the result of algebraic computations reducing the problem of finding the mentioned direction to a standard mathematical problem, the minimization of a quadratic form, $\min k^T S k$. The information in the ST and LST are fully equivalent (one to one, onto) and as follows

$$I_{20} = (\lambda_{max} - \lambda_{min}) e^{2\angle u_{max}}$$

$$I_{11} = \lambda_{max} + \lambda_{min}$$

where the largest eigenvalue $\lambda_{max}$ defines the corresponding eigenvector associated with it $u_{max}$, the structure tensor. The Difference between the eigenvalues express dominance of the fitted direction $u_{max}$ as compared to its orthogonal, $u_{min}$. This difference is an evidence for the presence of a unique orientation,
3.3. DENSE VECTOR MAPS AND INTERACTIONS WITH EXPERTS

albeit it is energy (contrast of an edge) dependent. Absence of linear symmetry assumes randomly oriented gradients in the support (defined by the outer scale $\sigma_{out}$ of ST), or occurrence of (two, three, four...etc) jointly occurring and maximally different orientations therein.

The inequality $|I_{20}| \leq I_{11}$ holds with equality if and only if a single orientation fits the image perfectly, in which case $\lambda_{\text{min}}$, the Total Least Square (TLS) error of the fitting in the Fourier domain, vanishes, [1]. This is a case of unique orientation and eigenvalues being different, i.e. $\lambda_{\text{max}} >> 0$ and $\lambda_{\text{min}} = 0$. In other words, all local gradient vectors (in double angle representation) are pointing in the same direction. A vectorial summation (or average) of such (parallel) vectors then equals to the arithmetic (or average) summation of the magnitudes of the vectors. The component $I_{11}$ is used to normalize the complex $I_{20}$, the magnitude of which becomes a quality measure between 0 and 1, so that the magnitude of the result specifies the quality of the estimated orientation in the support of ST, and the argument specifies the estimated orientation.

The complex images $I_{20}$ or $I_{20}/I_{11}$ can then be directly interpreted by humans since they can be displayed as color-images, by steering H and V components of color pixels, represented in HSV color space. Therein, argument, and magnitude of complex pixels would steer H (scale) and V (intensity) respectively, at maximum saturation. Having two degrees of freedom, the normalized complex pixel values (subpart of the structure tensor) allow it to be represented by two color components of a (three dimensional) color space. In addition, HSV space complies with natural human colour perception, where brighter parts (higher intensity V) of at highest color saturation (=vivid colors) deservedly draw our attention. Orientation map, Fig. 3.8, is an extension of commonly used images with gray valued pixels to a higher pixel dimension. Pixels values are then (orientation)vecors in double angle representation, discussed next.

The local ridge orientation is an angle $\theta \in [0, 180]$, masking the difference in orientation between $\alpha$ and $180 + \alpha$. Full range of 360, called direction here next for convenience, will differ the ridge direction angles that are equivalent. Therefore, we need to use a (ridge) direction representation that is invariant to a rotation of 180 degrees, which, additionally allows for a continuous (smooth, without jumps) when ridges rotate. Such a mapping is achieved by doubling the angle of ridge gradient directions. However, this is already technically achieved by the argument of $I_{20}$ which corresponds to the double of the (common) unique direction angles of parallel ridges in the support of ST.

The support of an ST is illustrated in 3.9 Right. From an engineering point of view $I_{20}$ is nothing but a vectorial averaging of (squared gradient) vectors in such a support (not shown) except that the vectors are internally in double angle representations before entering the vectorial averaging. It is thus an average of "rectified" gradient vectors, such that those that differ with 180 degrees no longer do so after rectification, i.e. they are equivalent (and do not
sum up to zero). The quantity $I_{20}/I_{11}$ is a vector where the orientation angle is present in the direction of the vector in double angle. The quality of the angle estimation is reflected in the magnitude of the vector. The angle representation is however different than the commonly used gradient image representation in color where there is a direct (not two-to-one) one-to-one connection between the ridge (orthogonal) gradient angle and a hue representation of that angle. Consequently, it is all the more important to have a reference circular image, Fig. 3.8, acting as a legend for what a forensic expert not habituated to, the color/orientation representation he sees.

To use a metaphor, in the era of black and white movies we had no possibility to be informed on the color of objects in movies. We were certainly aware of this information loss when we were watching the movies of Charlie Chaplin but we had to be content with what was available. Likewise, the forensic experts are aware of that there is actually more to a fingermark than the constellation of singularity points, but this is all what experts are offered to communicate with machines. Machines can currently only extract and match with this information in a way that an expert wants it. However, the expert wants also to use more image information and pass it on the top of key points but is hindered by that the machine is often not able to do the same on tenprints and certainly not in fingermarks. We have to remember that already extracting the minutiae from fingermarks is generally a huge challenge for machines (although the task is feasible for them in tenprints). The contrary is also a problem because what machines can do is too difficult and/or meaningless for humans who have to deal with its consequences, e.g. experts, judges, attorneys, jury, defendants, and victims.

In this context we have suggested dense orientation and frequency maps even for fingermarks. We have suggested to represent them in color maps so
that humans can evaluate and correct them by means of GUIs and legends in an attempt to effectively improve the result of machines. In Paper [C], we have suggested to compact this human interpretable information and represent it as vectors to complement key points (just as the currently existing minutiae directions complement minutia location information) so that machines can use even dense maps when identifying key points. For this purpose, and the purpose of ridge counting, it is important that what is to be used by machines in subsequent automatic processing is extracted as reliably as possible. Although the machine can do most of the chores when extracting fingermark orientations (and frequencies) densely, the human expert must be offered to verify and correct the results when needed. The human correction must also be allowed to be "no color" for a region, which is possible in our representation. This would mean that it is not possible to extract the desired information with the currently available data, although the machine suggests otherwise for any reason.

When suggesting human experts to interact with machines via dense color maps, we have been assuming that it is easier for a forensic examiner to correct entire regions by delimiting them or brushing them with color semi-automatically. A motivation for this belief is that tracing the ridges one by one, as it is done currently, is error prone, not the least because of the fatigue, since the task is difficult and tedious. A reason is that backgrounds of fingermark images are often strongly structured and can interfere with fingermark structures.

3.4 Discussion

Commonly high density frequency maps and orientation fields are not calculated in fingerprint recognition (tenprint to tenprint) in published studies. Due to locally slow changes of ridge orientation and frequency, block-wise field estimations are enough to enhance good quality tenprint images. In case of matching fingermark to tenprint there is not nearly as much image knowledge available. Even if the enhancement is done on tenprints in the fingermark/tenprint combination, the process of matching may need the full information of tenprints in tact to achieve a better recognition. The evaluation of the frequency map of [22] (low-resolution dense map) against ours (high-resolution dense map) evidences a measurable difference in favor of high resolution frequency maps as detailed in paper [B].

The frequency estimation algorithm was tested on overlapping fingerprints to demonstrate its use in multiple plane waves. The SD27 fingermarks contain numerous minutia neighborhoods containing multiple planar waves, violating single frequency and/or single orientation model. Experiments show that for such cases other enhancement algorithms, e.g. [32], do not offer more discrimination power to identity establishment. We think that extracting a single
scalar from Short-Time Fourier Transform mathematical model (the position of a peak) as frequency estimation does not handle more joint occurrences of multiple wave patterns than the suggested algorithm if no additional heuristics are utilized.

The current implementation of an orientation map is done by computing STs, the parameters of which are obtained from the (global) average frequency of the companion frequency map. In this (iterative) process it is assumed that fingermark orientations have slow variations, and therefore ST parameters are set globally, as determined automatically by the average frequency. Although the (dense) frequency map is not used directly in the orientation map, it is used indirectly. The enhancement of the original is done by projections on automatic Gabor filter constructions for every pixel, which demand dense estimations of both orientations and frequencies. Accordingly, dense frequency maps are implicitly used to construct dense orientation maps and vice-versa, via the enhanced image. This makes sense because fingermark areas left in crime scenes are small (in comparison to tenprints) and variations in orientations can be expected to be slow.

Experiments show that single parameter estimation for all STs in a fingermark was sufficient to yield favorable results in several tasks of forensic identification, i.e. ridge counting, using frequency as minutia descriptor, in comparison to state-of-the-art. However, further improvement can still be envisaged as future development of the method. For example, estimating orientation and frequency maps can be extended to include larger areas but one sub-window at a time, and assuring that sub-windows overlap for continuity.

Forensic routines attempt to improve the quality of an identification by providing expertise from two different examiners. It seems to be beneficial if this procedure will be additionally supported by machines acting as "calculators" on demand for fingermark/tenprint measurements. For example, ridge counting between all key points can be time consuming, since the number of minutia pairs grows quadratically with the key points. Such exhaustive counting may even not be desirable, even if it is easily computable since the results would be too cluttered and not all counts would have a high reliability. Using the results in Paper B, a machine could automatically select a subset of the key points (assuming that they are already extracted) and count the ridges between them. The presentation of the counts should show the selected pairs and where the estimated locations of the ridges are. The forensic expert would then be able to verify the results quickly and accept, reject, or even redraw paths (not necessarily straight lines) between pairs of key points increasing the total reliability of her decision making. The machine would not make a decision but merely support the human expert with easily verifiable (by appropriate GUIs) computations.

Ridge counting ground truth was not available in SD27. We performed instead a manual verification for 100 minutiae pairs since exhaustive verification would be prohibitive (on +70,000 paths). We counted then the corresponding
paths in both mates (tenprint/fingermark) visually. The manual verification suggests accurate computation of ridge-counts in conformance with the reality. Since the locations and the labels of minutiae of SD27 were available from Paper A, the two ridge counts on the same path in both mates should also agree, if the counts and locations are accurate. This allowed an additional (automatic) evaluation of the ridge counts. As detailed in the paper, between 78 % and 82 % agreement (within 1 ridge distance error) could be verified using +6000 through +70000 different paths. Remembering that the disagreements could also be due to inaccuracies in localization of the minutiae, and that a human expert can verify the computed ridge locations by a graphical interface, the results support the conclusion that our ridge-counting is accurate on fingermarks.
Chapter 4

Extracting features from dense vector fields

4.1 Motivation

"Tärta förstås", sa Lillebror. "Jag får en födelsedagstärta med åtta ljus på."
"Jaså", sa Karlsson. "Du, jag har ett förslag!"
"Vad då", frågade Lillebror.
"Kan du inte be din mamma att du får åtta tårtor och ett ljus istället?"

Karlsson på taket, Astrid Lindgren

Can a person be identified with characteristics of the image around a single point (minutia)? Put differently, can a point be characterized by the image information around it so that the point becomes so unique that it can be recognized among other points? If the latter is not strictly achieved, i.e. can the point be characterized to belong to a unique category of fingers, instead of a unique finger? These were some of the questions that have motivated the study behind Paper [C]. They are also relevant to other applications of image analysis, thus fundamental, we believe.

We suggest novel descriptors representing image information around key points to enforce the uniqueness of their "identity". Thereby we wish to enable a machine algorithm to pull out the matching tenprint of fingerprintmark based on fewer minutia but with extended descriptors. Each of our descriptors are equipped with its own quality measure. Sample fingerprints with as few as 5 minutiae were possible to establish identities for, e.g. Fig. 4.1. An additional reason for why image based novel features are needed in forensic fingerprint recognition is that they occupy small (physical) areas compared to tenprints. Accordingly, as much information as possible should be extracted from them, not only singularity points. Current state-of-the-art systems used by forensic
experts employ coordinates of singularity points and their directions only. This is often not enough for identification due to the too few available singularity points. We propose usage of rotation invariant features since the global orientation of fingermarks (the orientation of the finger as it leaves a fingermark on a surface) found on crime scenes is often not possible to estimate accurately. Also using rotation invariant orientation features allows us to make each minutia independent of other minutia, i.e. minutia constellation.

Additionally, the local structure of the fingerprint is less stable as compared to the global structure. Even if a fingerprint quality is poor, matching is still possible if entire fingerprint available is large [33]. Forensic fingerprints often possess less amount of image and minutia information, as well as complicated noise types, constraining us from using and comparing to the methods ([34], [14]) working well for traditional (biometrics) fingerprint images.

4.2 Approach Description

"Think globally, act locally"

As a feature to describe the local image structure around a key point we use orientation or a change of gray value in one direction. To be more precise we describe the dense orientation field around the point rather than the image. Later, the concept is extended to describe dense frequency maps, because as it has been shown in Paper [B], even the local frequency is an angle, at least as it is delivered by the estimation algorithm. When more advanced structured than constant orientation are allowed in the model, the model can extend the simple orientation concept described by LST (3.3) to include other isocurves than straight lines.

The Generalized Structure Tensor, GST, can be defined in curvilinear \((\xi(x, y), \eta(x, y))\) coordinates, where \((\xi, \eta)\) constitute a harmonic function pair. GST has the ability to detect the orientation of isocurves of a local image, when orientation is interpreted in harmonic coordinates \(\xi, \eta\). An easy way of obtaining harmonic function pairs is use of analytic functions as generators. The real and imaginary parts of an analytic function, \(g(z)\), are always harmonic pairs, i.e. \(\Re(g(z)) = \xi(x, y), \Im(g(z)) = \eta(x, y)\). The ordinary ST is obtained thereby as a special case of GST by the simplest analytic function transforming \(z\), namely the identity transformation, \(g(z) = z\), mapping \(z\) on itself, where \(\xi = x, \text{ and } \eta = y\). The ordinary ST measures the orientation angle (doubled) of isocurves of the type \(ax + by = constant\), i.e. the \(2 \arctan(b/a)\). GST measures the orientation angle (doubled) of isocurves of the type \(a\xi + b\eta = constant\), i.e. the \(2 \arctan(b/a)\). In other coordinate systems than Cartesian, the isocurves (the orientation angles of which are measured) are not straight lines. When such isocurves are "parallel", it is meant that they are parallel in the \((\xi, \eta)\) coordinates, i.e. they share the same \(2 \arctan(b/a)\) but each such curve has its own \(constant\) in \(a\xi + b\eta = constant\).
4.2. APPROACH DESCRIPTION

Figure 4.1: Sample fingerprint of SD27 database with 5 minutia points used for identification (pink). Additional three points were extracted by the examiner at the analysis step but later excluded as not present in the corresponding tenprint (green). Thus there are two minutia annotations available, one called Matched set, where every minutia in a fingerprint has a matching minutia in its corresponding tenprint, and the other called Ideal set, where in addition to common minutiae there can also be minutiae that do not exist in the matching fingerprint. These sets are detailed in Chapter 2 and Paper [A].

For example, isocurves of a "spiral" space are defined by the following equation

\[ g(z) = \log(|x + iy|) + i \arg(x + iy) \]  \hspace{1cm} (4.1)

Members of spiral families have built-in rotation invariance which is a desirable property in forensic applications where the global (finger) orientation often is not known. In GST one filter detects all members of the modeled isocurve family. This applies to the spiral family too, i.e. the presence of any spiral
family member and its orientation angle are estimated by a single filter devoted
to the spiral family. The orientation angle of a spiral is its level of twist or
chirality, see Fig. 4.2 where these angles are marked for two spiral images.

The list of symmetric patterns detectable by GST can be made infinitely
long. This is because any analytic function \( g(z) \) generates a harmonic curve
(coordinate) pair \( \xi, \eta \), and all patterns having iso-curves \( a\xi + b\eta = \text{constant} \)
will be detectable provided that \( \xi \) (or \( \eta \)) is known to GST. This detection
of symmetry patterns and measure of the angle \( 2 \arctan(b/a) \), is equivalent
to fitting a line in the Fourier domain of the image containing the pattern
where the Fourier Transform is taken w.r.t. \( \xi, \eta \) coordinates (not \( x, y \)). However, ST as well as GST is able to detect the presence of the patterns and
estimate their "orientation" (\( -2 \arctan(b/a) \)) without Fourier transformation,
directly in spatial domain. For this GST uses dense orientation fields, i.e. those
generated by ST, to detect its symmetric curves and estimate the orientation.

In Paper [C], we have suggested to use a particular set of analytic functions,
Table 4.1, to generate harmonic functions and the series of symmetric patterns
associated with them, as illustrated by top row of Fig. 3 of the paper and
generated by the (real and imaginary parts of the) analytic functions.

Each of the real and imaginary parts of these functions represent a curvi-
linear coordinate system in which a "line" is actually a curve in Cartesian
coordinates. If an image around a point has isocurves as described by one
family of these curvilinear coordinates then the local image has an orientation
can be detected by a GST specific for that curvilinear coordinate system. The
detection of the orientation is done by applying a single complex filter, de-
termined by the analytic function above, Paper [C] eq. (8), to the (ordinary)
orientation map of the original image (\( I_{20} \), or \( I_{20}/I_{11} \), as suggested in paper
[B] for fingermarks). These complex filters, turn out to have integer angular
frequencies in the complex term, Paper [C], eq. (8). From this it follows that
the above analytic functions are not randomly obtained but these are selected
such that their corresponding filters constitute a Fourier basis in the angular
direction. As such estimating the orientation in an image (neighborhood) in
the above series of curvilinear coordinates corresponds to analysis (by recon-
struction) of the dense orientation map via the orthogonal Fourier basis in the
angular direction.

Furthermore, the analysis (by reconstruction) is extended to the radial
direction where the center is the common (singularity) point by suggesting a
novel way of obtaining GST filters where the common point is the origin of
4.2. APPROACH DESCRIPTION

Figure 4.2: First column (left) is a ring around a fingerprint and its orientation map in color. Next column is the parabolic curve family having zero degrees orientation in a ring with the complex GST filter detecting all members of the parabolic curve family (color ring). Different members differ with global rotation of the zero degrees pattern. Third column shows two members of the (same) parabolic curve family with the intrinsic orientation of each pattern shown in yellow. The fourth column is ditto but showing two members of the spiral family, with their respective orientation angle marked with yellow. The orientation angle of a member of a harmonic curve family (e.g. those in the third and fourth columns) is determined by the orientation of a line in the respective curvilinear coordinates. For example it is $\arctan\left(\frac{b}{a}\right)$ where $a, b$ are coefficients in the equation of a spiral $a \log r + b\varphi = \text{constant}$, marked as the angles in the Fourth column. These intrinsic angles (yellow) are measured by GST along with a quality quantity representing the goodness of curve fitting.

In previous studies, in addition to orientation maps, reinterpretations of the linear structure tensor have been used to find interest points, in 2D [35] as well as in 3D [36]. The idea is based on detecting lack of a dominant orientation. The concept came to define "cornerness" and it has been used in
many applications. This is because a presence of orientation as well as a lack of orientation is verifiable by use of ST. This is typically done by heuristically constructed algebraic expressions testing the closeness of the eigenvalues of ST. Harris corner detector, [37]

\[ \lambda_{max}\lambda_{min} - \kappa(\lambda_{max} + \lambda_{min})^2 = (\lambda_{max} - \lambda_{min})^2 + \kappa^{-1}\lambda_{max}\lambda_{min}, \quad \kappa \approx 0.04 \]

(4.2)
is an example where the lack of orientation is tested rather than presence of a corner because there is no explicit corner model in their study. Note that this algebraic expression of eigenvalues becomes large when the eigenvalues in the first term are close to each other, if at the same time both eigenvalues are large causing the second term to surge signaling the presence of corner (by way of lack of orientation). Forstner and Gulch [38] have taken a similar approach to quantify cornerness, via an algebraic testing of the closeness of the eigenvalues of ST.

There is a concept of "Tensor" processing, [39], [40] that is related to ST, [1, 41], and used to apply to images with high dimensional coordinates, \( x_1, \cdots x_n \) where \( n > 2 \) and the coordinates are Cartesian. Example image types include sequence of images as in video or slices of a volume as in computer tomography. The tensor is computed for a local image around an image point and is capable to tell if the neighborhood belongs to one of the \( n \) predefined categories. In \( n = 3 \) these types are called "plate", "stick" and "ball", [40], and are easily found by decomposing ST (eigenvalue decomposition) into a sum of plate, stick and ball tensors. The decomposition (and thereby classification and the subsequent processing) is always possible given a ST of a point, because ST encodes the statistics of the concentration of energy in the Fourier Transformation of the local image, [41], [21]. By way of example there are 3 types of extreme concentrations, plane, line and point, in 3D corresponding to 2-d, 1-d, and 0-d hyperplanes on which energy concentrations can occur in the Fourier domain, etc. Thus this extension of Tensor concept is purely in the number of Cartesian dimension, \( n \), and seeks to explain the local image by the ordinary ST albeit in n-D. An interest of this decomposition in terms of the natural concentrations of hyperplanes (in Fourier Domain) is that the tensor decomposition automatically suggests algebraic expressions for testing presence of planes, lines, and points (=corners) by orientation statistics. ST decomposition into its more basic tensors (plane, line, point) brings a systematization to interpretation of the eigenvalues of ST, without heuristics.

The GSTs, which we have used in Paper [C], are applied uniquely to images with coordinate dimension \( n = 2 \), where there are only two types of Fourier Transform concentrations, line and point concentrations. However, it is the coordinate types that are generalized in an infinite number of variations, not the number of dimensions. Thus not only Cartesian coordinates, \( x, y \) but also each harmonic mapping of them as obtained by an arbitrary analytic function \( g(z) = \xi(x, y) + i\eta(x, y), \) e.g. \( \log(z), \sqrt{z}, z^2, \) etc. (with \( z = x + iy \)) can gen-
erate a GST which describes the statistics of concentration of the FT (taken in the respective $\xi, \eta$ coordinates). For a taxonomy of tensors and the orientation types associated with them we refer to [21].

In applications of fingerprints, two sets of harmonic families have been used in GST to detect singularity points of type cores or deltas [42]. These were used to quantify if the entire neighborhood, not just a ring, can be explained by members of harmonic families. In paper [C], we have used these (and more) families too, but to describe (image) rings around a point of interest (rather than the entire neighborhoods).

We have suggested to use the ability of GST to detect the presence of curve families and to measure their orientation angles around an arbitrary image point. We have applied this to forensic fingerprints and iris images to establish the human identity behind such images. In our approach we use first the ordinary structure tensor to obtain the orientation maps, then subsequently we apply the generalized structure tensor to obtain feature vectors from these maps. This establishes a mathematical framework for extraction of visually meaningful features, and at the same time quality measures associated with them via GSTs, which we believe is useful for also other image analysis applications. Implementation of GST is a discrete convolution between LST of an image and a (novel) series of filter family similar to symmetry derivative filters angularly, [20], but more precisely steerable in the radial direction, see paper [C] for details. The convolution kernel (filter) is complex valued, as are the symmetry derivative filters, resulting in complex valued feature vectors. We represent the complex filters as color images, using the same displaying technique presented previously (Sec. 3.3) in Fig. 4.2 Left.

Details of the algorithm of feature extraction are described in the paper. However, a stylistic version of it, and the context it is used in, is illustrated in Fig. 4.3. In a preprocessing step, an enhanced orientation map of the original fingermark is obtained iteratively as detailed in Paper [B]. The enhanced orientation map is displayed as a colour image. The expert is given a chance to correct the regions with false colors (orientations are estimated but they do not correspond to those of fingermark ridges) or no-colors (poorly estimated orientations). The expert manually or semi-automatically extracts a set of singularity points. Then these are presented to our feature extraction algorithm. The algorithm decomposes the orientation map within concentric rings around each singularity point into a set of predetermined family of curves. Concentric rings are obtained by means of filters with radial functions having peaks and peak-widths progressing geometrically, to preserve and use the available information around singularity points systematically. We have developed a framework for controlling the overlapping areas of filter functions and the locations of the peaks by few parameters, as summarized in Lemmas of the paper [C].

We demonstrated that our feature extraction generalizes to describe other vector maps well, not only those representing local curve orientations. In paper [B] it was shown that even spatial frequency information and the quality are
CHAPTER 4. FEATURES OF DENSE VECTOR FIELDS

Figure 4.3: Illustration of the context of our feature extraction method and the summary of the features

estimated as angles and scalars, respectively, that can be merged into a vector field. This was possible with the discovered connection between the frequency and the orientation of a line fitted to the log-space of the trace of the ST. In Paper [C] we have shown that the frequency map around singularity points contains valuable, identity specific information, that can be extracted and encoded by our features. This was observed already in Paper [B], where the frequency map as is, without compaction by SAFE features, was shown to contain valuable (minutia) identity specific information.

Both an orientation map and a frequency map originate from poor quality images in forensics. Accordingly they cannot always be reliably extracted with a lights-off algorithm, including our iterative algorithm presented in Paper [B]. However, using a color map, the examiners can be equipped with the ability to correct semi-automatically the automatically extracted orientation and frequency maps by painting in them with user friendly GUIs. Our feature extraction in paper [C], has however been applied to fully automatically extracted orientation and frequency maps, to allow evaluation of its results.

As mentioned above, we have proposed to project the orientation map of an image at different tori around points of interest on a base of harmonic
functions. These projections, themselves are structure tensor computations but in the coordinates defined by curvilinear bases, Fig. 4.2 Left. This procedure has allowed us to construct a feature vector for each key point, e.g. minutiae, of a fingerprint image. Such a projection can be seen as a verification of the conformance of the input data with a model (as represented by a filter, which in turn represents a curve family), making our feature extraction model-based. Model-based feature extractor provides grounds not only for unique description of a signal (by means of the model), but also for an error (certainty/quality) estimation of a fit. While we project the complex valued orientation map on the complex filter, we are thereby searching for a specific symmetric structure in the underlying gray images, Fig. 4.2. Such a projection result is a complex scalar which is the filter response, as all scalar products are implemented via convolutions.

We have shown that the filter response magnitude represents a quality measure for the presence of a member of the curve family, in the corresponding torus. At the same time the filter responses are coefficients of projections on orthogonal functions. This double meaning of the coefficients allows to conclude, by way of example, that the ring neighborhood of point A (Fig. 4.2 Left) has 0.1 amount of parabolic symmetry, 0.8 linear symmetry and 0.01 hyperbolic symmetry and 0.05 amount of circular/spiral symmetry, etc., which is description of the content but also allows reconstruction. The latter is because the orientation map in the ring can be reconstructed by these coefficients. The argument of each complex coefficient (filter response) tells which family member of the corresponding curve family best explains the torus, since for each angle there is a unique "orientation", though the meaning of "orientation" varies from curve family to family, e.g. Fig. 4.2. Because each filter response is complex valued, it can be displayed as a color image, using the same technique as (HSV color space) described in Sec. 3.3. In a way, one can say that our features are more "colorful" than state-of-the-art because they are orientations with well defined visual interpretations (in terms of curve-families) and built in quality measures (in terms of Total Least Square Error fits).

Our feature vector represents the orientation map around minutia compactly (with tens of complex scalars we are describing hundreds of complex pixels in tori), and it contains both dependent and independent components of global (finger) direction. Absolutely rotation independent components are obtained by projecting the ordinary orientation map (of lines) on the (orientation map of) Spirals Space. The global direction steerable/dependent components are obtained similarly but by projecting the orientation maps of tori on other curve families, (achieved by complex convolution). The global direction dependent filter responses have however an interesting property namely that when the image (finger) rotates globally with \( \Delta \varphi \), the filter responses arguments get translated with \( p \Delta \varphi \) with \( p \) being an integer determined by the symmetry order of the curve family (of the filter). That is the filter responses are (mathematically) invariant to rotation in the sense that the filter responses
Figure 4.4: Illustration of two SAFE feature vectors as two images before and after rotation compensation with a known global rotation. Left: the features are before compensation, Right: the features are after compensation. Note that the green column corresponds to the spiral family which is invariant to rotation. Each row corresponds to a torus whereas a column represents the same symmetry (curve family).

are compensatable to known rotations without a need to project the rotated images. In paper [C] this mathematical invariance is detailed in Sec. 3-D, and in eq. (18). Here Fig. 4.4 illustrates the rotation compensation of presented features.

One can notice that all the filters are highly symmetrical and the underlying gray-image patterns are also highly symmetrical, explaining the acronym we chose for our descriptor, SAFE (Symmetry Assessment by Finite Extension).

Independent of ridge counting, and dense absolute frequency maps, orientation maps are critically important for automatic fingerprint processing e.g. for minutia, core and delta points extraction. It is also typical for fingerprints that the quality of ridges in certain regions of the image are so poor that the ridges in such regions are only perceivable thanks to the (somewhat) higher quality in the neighboring regions. Built-in quality estimates exist on the two critically important layers of signal processing of our features. First, they help to reveal the amount of the available high quality data even before orientation maps are input the feature extraction processing of a fingerprint so that a human expert can verify and correct them. Second, the explanation power of each symmetry type for a particular (ring) region is clearly visible if the feature vector (which is a matrix) is displayed as a color image, e.g. Fig. 4.4.
4.2. APPROACH DESCRIPTION

The resulting feature vector can be used to improve the odds of extracting the matching image from the database of tenprints. Our and a previous study [43] show that angle and coordinate of minutia as feature, can be complemented with image based information to improve the performance and the quality of the fingermark identification. Possible additional features can include ridge counts, which are done manually by human examiners currently, but can be transferred to machines which can count and present them with appropriate GUIs for verification by human experts.

All image information around the point of interest must ideally be taken into account, which is important for forensic images where only small amounts of useful information are available. The suggested mathematical framework includes finite expansion of the orientation map of an image, where each projection is an independent GST in a well controlled torus. This framework facilitates to adapt the method to different applications because the meaning of each feature is well understood in terms of (a priori known) symmetric curve families.

All curve families but one (which is the curve family representing straight lines, ridges and edges, etc) which we use in GSTs are well localized, i.e. object-based. Each such curve family has a unique (mathematical) singularity at the origin of the (curvilinear) coordinate system. This fundamental singularity existed in the original filters of the respective GSTs detecting the respective curves, [20, 44, 45], and are evidently present in the novel filters too. Paper [C]. It means that the location of a point of interest matters w.r.t. the surrounding image pattern and vice versa i.e., the extracted features would change significantly if the feature is extracted from even a point nearby the true singularity point. This behaviour is desirable since the features attempt then to extract point specific rather than region specific image structures. Consequently, the feature vector will attempt to equip the point on which it is associated with, with a unique identity extracted from the image pattern, hopefully, present in the neighborhood. If this information is not present in the neighborhood, the result will be a complex value close to zero.

In the paper, we have therefore put forward that the SAFE features, as descriptors of points of interest, measure the particle nature of the surrounding image neighborhood. This is to be contrasted to what we called wave (textures), e.g. linear symmetry features. Wave features should be constant when extracted for local neighborhoods of a wave, meaning that they should not at all be sensitive to location if they are extracted in a completely different location of the image, as long as the new location is part of the (same) wave. It is a fundamental nature of waves that they are repetitive and everywhere, (not localized).

To be precise, features measuring the wave nature are present even in a SAFE feature vector, one per torus, wherein we have extract ~10 times more features (the linear symmetry component). If an application demands it however, it can be deleted from the SAFE vector, since its wave nature, as well as
which filters they correspond to (and therefore not to be used), are established. We have included it deliberately for two reasons. First, it is not sure that it is desirable to ignore all wave type characterization of a neighborhood. An expert may want to include information sensitive to whether or not a minutia is close to a core or a delta point in the information describing the minutia. If the minutia is peripheral (w.r.t. cores or deltas) then it is known that it tends to have little orientation variation around it. Linear symmetry measures exactly this, the presence of a uniform orientation. Second, The collection of GSTs become (mathematically) complete to represent any non-zero orientation map (per torus), since this GST represents the DC component of the angular Fourier Transform. It means that if the norm of a SAFE vector equals to the norm of the DC component, then there is nothing but a planar wave with a perfect orientation (a sinusoid in the torus). This information can be or not desirable to know in an application. The analogical metaphor that an application may reject or accept a visual feature is 6 and 9 in digit recognition. In many applications we want to measure rotation invariant visual properties (including fingerprint recognition), but in this case we do not, else we can not differentiate a 6 from a 9, etc.

As mentioned, Linear symmetry features which correspond to curves having a common (constant) direction, are also detectable by GST formalism. The GST reduces then to an ordinary ST, with its filter (outer scale) detecting the curves being a Gaussian in the original formulation, [1]. A Gaussian is a real valued function, as opposed to filters detecting the other symmetric patterns, which are complex. In other words neither the curves, nor the filter detecting linear symmetry have neither intrinsic, nor unique (singularity) point to anchor the neighboring patterns to. Nonetheless, in SAFE features, the associated filter of the Linearly Symmetric patterns is a real torus having a singularity at the origin.

This singularity is not an essential singularity of the Linearly Symmetric patterns that the filter aims to detect. The singularity in the corresponding SAFE features is induced by our desire to define the (spatial) support of the (G)ST, to be a torus, which in turn is caused by our desire to make the SAFE features complete. However artificial or external it may appear, this "imposed" requirement has actually never been contradicted by the theory of ST. This is because the Gaussian filter in ST, even in the original formulation, was meant to define the (spatial) support in which the Linear Symmetry features were to be measured. That ST has been used to extract local features in circular (filled, not rings) neighborhoods, does not exclude its ability to model and measure the same but in any other type of spatial support, (e.g. a ring, two attaching rings like the digit 8, etc).

Because the convolution of a Gaussian with a Gaussian is another Gaussian, and the convolution of a Gaussian with a (complex) torus is another (complex) torus, the complex filters of GST originally suggested to be applied directly to Infinitesimal Linear Symmetry images, i.e. to complex images of \((f_x + if_y)^2 =\)
4.2. APPROACH DESCRIPTION

\[ f_x^2 - f_y^2 + i f_x f_y \] (without smoothing by the outer scale Gaussian of the ST), can be applied to any orientation map represented in double angles (of gradient directions) and the responses still retain the same meaning (of fitting members of curve families to the image). This is what we also did in Paper [C], i.e the input to the finite expansion of SAFE features is an (enhanced) orientation map, i.e. STs computed at local circular supports.

Accordingly, the intention behind using linear symmetry in the "opposite" sense to its nature, which is to identify the presence of linearly symmetric regions, is thereby justified and is not opposite to the original meaning of ST. This is because we use them to explain global variations of orientation maps around minutia points. We are aware of that they are translation invariant (wave nature) even when used with real valued torus filters, which do have a singularity. This use of Linear Symmetry extraction (orientation) in tori, is however, in a much coarser level than its other use to extract the orientation maps, paper [B], where the scale is finer (at most a few ridge distances). In both cases however the support (ring and Gaussian respectively) are purely real (else it is not linear symmetry orientations that are measured per construction. We recall that only when the spatial support of the filter is complex when filtering \( I_{20} \), we obtain fitting of non-linear curve structures. This was used to detect singularity points such as cores and deltas of fingerprint images [42], which have even larger spatial supports our tori.

A SAFE vector at an image point is intended to uniquely define the point (e.g. a minutia), via the vector field patterns around it. We have used complex filters (nine per tours) which were originally suggested to extract (orientation) parameters of curvilinear patterns at exponentially increasing radii, in tori with exponentially increasing thicknesses to describe the true orientation maps (vector fields) but also the frequency (represented as vector fields, Paper [B]). We presented experimental results of identification and verification of single core points and minutia points in fingermarks against their tenprint correspondences, by using our SAFE features on both types of maps, true orientation maps, and artificial orientation maps (frequency maps). The latter was possible because SAFE features, being finite expansions of vector fields on one hand are not limited to represent orientation maps as they come from STs. On the other hand, we can no longer interpret the delivered coefficients (elements of a SAFE vector) as evidence for membership in harmonic curve families. If a coefficient magnitude is large we cannot say that it represents an orientation singling out one member of the harmonic curve families detailed here and in the paper, because the input vector field is not a field representing the local orientation, (here the absolute frequency which happens to be in a vector representation).

The SAFE feature vector remains invariant under global rotation (when proper rotation compensation is made). Such a compensation can be made according to an externally provided orientation (e.g. the orientation of the
minutia delivered by a human expert) or , provided that it is non-zero, also automatically by the complex response of the parabolic symmetry in any torus.

All SAFE features are by definition also invariant to the true location of the singularity point (global translation invariance). That is the feature vector extracted at a point of a tenprint, and the one extracted in the corresponding point of the matching fingermark will be the same, if the fingermark has undergone a translation (only) w.r.t. the tenprint. This assumes that the neighborhoods of the respective points are fully visible to the feature extractor. However the SAFE vector will change if it is placed in a false location, i.e. not exactly at the same singularity point (local translation sensitivity). This desired sensitivity should not be confused with its desired insensitivity to global translations, because the latter insensitivity is when the true location is translated elsewhere in another image whereas the former is due to the translation of the current point (which is a true point) to another point (which is a false point, even if the displacement is small).

In the experimental section of paper [C] we have extracted SAFE features at minutiae and cores. The reason for not using delta points is their more complicated interpretation in terms of their three-folded ambiguity w.r.t. global rotation, Fig. 4.5, in case some delta points really turned out to be perfectly three folded ambiguous. This in turn would be less effective to demonstrate how our rotation compensation works, pedagogically. This does not mean that all delta points are perfectly three folded symmetric and are therefore useless SAFE features in fingermarks. By contrast, most individuals deltas are not perfectly three folded, meaning that there is a good chance that they can provide distinctive identity. Such a study is however left as future work.

We have used the annotation of SD27 when performing our experiments on recognizing cores and minutiae deliberately, to free SAFE features from the burden of i) having used locations that are not extracted by humans whereas we have suggested the features for exactly such a scenario, ii) having used location information that are not easily repeatable or accessible if automatic techniques would have been used for localization. However, automatic estimation of minutia and core directions (as well as their locations) is possible with the same GST, at least for tenprints, [46]. A possible study of SAFE features, do include GST’s, to suggest automatically the minutia and core points of a fingermark to a human expert is left as a future study.

The SD27 database comprises 258 tenprint/fingermark pairs while the number of cores existing both in tenprint/fingermark mates is 263. This is explained by that, some pairs of tenprint/fingermark mates had more than one core each (e.g. loop) and some pairs had none.

As it was established when sanitizing the database, Paper [A], and also observed (but not detailed) by [43], there are two fingermarks in SD27 having the same tenprint mate but with different areas of the latter matching the two fingermarks. Thus the same tenprint is present in two different tenprint/fingermark mates. If we would suggest identification on a person level,
4.3. DISCUSSION

one pair needs to be excluded, at least from impostor tests. However, in order to analyze the identification power of a single point, we facilitated the point level identification and therefore left both images. It is worth noting that fingerprints are originated from (almost) non-overlapping areas of ten-print. Common minutia were excluded from the experiments.

We have presented recognition of people by their periocular regions of their faces. The periocular dataset we have used (BIOSEC database) has required less annotation expertise and therefore provided a larger test set in terms of the number of (periocular) images used in the evaluation (1200), although the number of different people was (75) less than that of SD27. In the identity recognition experiments we have sought support for the usefulness of the proposed SAFE descriptors in a problem where we knew that i) Gabor (magnitude) were among the most successful feature extractors for the problem, and ii) that Gabor (magnitude) responses excel in texture description, i.e. equipping regions with identity. The results confirm that SAFE descriptors nonetheless succeeded to improve the recognition performance offered by Gabor features when used in combination. However, they do not outperform the Gabor features when both used alone. We think that, this is because the iris region has both texture (wave) and object (particle) type of information although texture type information is the most dominant. This supports the view that SAFE features extract (object type) information that is not accessible to powerful texture features, because they are not as good object type feature extractors.

4.3 Discussion

In the paper we have suggested a set of features, SAFE descriptors, that extract image information around key points and can bring in novel recognition power to the traditional features, (minutia, core and delta points, and their directions).

SAFE features are orientation steerable, i.e. there is no need to rotate the reference image towards the test image, since the descriptors can be compensated for rotation by complex multiplication. In that sense they are rotation invariant. The spiral features subset is in any case absolutely invariant—no need to even rotation compensation.

We have provided experimental support for their usefulness based on single key point recognition. The end goal of these descriptors is to be complementary to the state-of-the-art features used for forensic fingerprints, a constellation of the key points. However, to extend the features such that their recognition power is combined with the existing key point constellation matchers is left out as a future study, as we had to focus our study.

The focus of our study has been to demonstrate the usefulness of image information, in particular as expressed by orientation and frequency maps.
The problem was to find a suitable mathematical framework to extract this information compactly and yet make sure that the extracted information is interpretable by humans. The latter was important, because the idea was to use a human expert to evaluate and correct the extracted features.

Accordingly our experimental and theoretical contributions had to focus on empowering single points with their own identity to the extent their individual image neighborhoods permit. This focus is also useful for in depth understanding of these features alone, isolated from effects of other modules in a full system of recognition procedure (which may or may not involve automatic machine processing). We think that this promotes repeatability.

The GST interpretation of complex filtering as an alternative to Generalized Hough Transform was introduced in [47] which in turn used the theoretical results that i) all analytic functions generate harmonic functions defining lines in curvilinear coordinates, and ii) the orientations of "lines" in such curvilinear coordinates can be determined using a single complex filter as suggested in, [45]. Three studies have been built on these studies, [48], [49], and [20] with partial overlap. The separable filters estimating $I_{20}$, given in [20] are similar to those suggested by [48] to describe the orientation in the local image.

The main novelties of [20] in comparison to [48] are i) it enhances $I_{20}$ that encodes the orientation of the pattern with the total error measurement $I_{11}$ so that these are fully equivalent to GST, ii) through the tensor, [20] fits a harmonic curve family to the iso-curves of the image that satisfies the total least squares optimality criterion, and iii) theorems on and introduction of symmetry derivatives of Gaussians, facilitating understanding the relationship of the filters (and the curve families) to each other. Paper [C] positions itself w.r.t. [20] as the introducer of a viable mechanism to steer the complex filters such that they are in geometric progression in their spatial support peaks as well as widths. Moreover, the idea of completeness in the radial direction is enabled so that GST responses, that can be extracted individually using standalone curve models (even without the function expansion idea), can be combined to constitute a finite (orthogonal) expansion allowing visually meaningful analysis (by synthesis).

In [50] it was shown that additional image based features have the potential to improve the recognition, but these results were due to (commercial) features which were not disclosed. We have here suggested novel image based descriptors whose inner workings have been analyzed theoretically as well as in practice. We are not aware of a prior published study using image features (to add-on to minutia features) in forensic fingerprint matching in general, on the only publicly available forensic database, SD27, in particular.

The SD27 contains fingermarks representing genuine orientation of ridges blended with background orientation. We think that the reported results are significant in that without using minutia constellation information, but by using only (automatically obtained) orientation information around key points (mostly one key point per latent) we were able to obtain a performance simi-
lar to minutia constellation. Evidently the problem of interfering background orientation is not resolved 100%. However, we think that the features contribute to this nonetheless, because an expert in the real scenario will be able to assess/correct the orientations around the few key points she/he deems important, before orientations are actually encoded into SAFE vectors.

This scenario (evaluating machine results) is realistic because in forensics, courts will not accept automatically extracted minutia from fingermarks. All minutia, core, delta points related measurements and other measurements from crime scenes must be verified by human experts. The neighborhoods of latents we used come from minutiae extracted by FBI experts whereby they are equivalents of case data. By contrast the description of such neighborhoods by our features were automatic in the experiments and therefore the results should be interpretable by humans to be verifiable by them. Therefore we think that the additional provision that the orientation (and frequency) fields can be verified/corrected by human experts because the fields can be visualized as color images is a useful novelty.

There is an additional provision that the machine can reconstruct an orientation map from a SAFE vector using the inverse Fourier Transform (of the SAFE coefficients), which can be presented to an expert to see what orientation map actually has been used in matching a minutia neighborhood in a fingerprint to that of a tenprint. This would increase the confidence of the human expert in her final decision. As this would involve a GUI development and it is difficult to evaluate it quantitatively, other than what we already did, we have left it out of the focus of our study, though useful to forensic examiners.

Testing and comparison to "other" automatic feature extraction algorithms was not available to us via prior published studies, since single point features have not been studied. Single point features that we have evaluated are not meant to be in competition with the key point constellation-based features, but meant to be used in completion to them. Our goal was to evaluate the usefulness of the automatically extracted image features from extremely poor quality images. The latter was evaluated by relating the image-based features to the benchmark, the constellation based identification, with the intention of fusion of both in the future.

The suggested descriptor evaluates the image information in the surrounding of a point. For fingerprint images if the point is a minutia then its neighbourhood is more or less expressive in terms of its orientation or frequency map depending on how far it is from a core or delta point. If the minutia is far away, then it is likely that the strongest symmetries in its SAFE vector are of the linear symmetry type (DC-components in tori). Evidently, such a minutia will only weakly be afforded a unique identity and a unique location using only the image information, Fig. 4.5 Left. If by contrast the minutia is close to a delta or core then other symmetry types (than Linear Symmetry) will also have significant magnitude. Such a minutia will then be highly likely to be described and located uniquely by SAFE features, Fig. 4.5 Right. This
CHAPTER 4. FEATURES OF DENSE VECTOR FIELDS

Figure 4.5: Two local images at different locations of a fingerprint, with corresponding orientation maps, *(Left column)* One far from a delta or core point, and *(Right column)* another in the vicinity of a delta point.

has motivated us to use minutia points close to deltas or cores in parts of the experiments in Paper [C], to study the object description capacity of SAFE features.

Nonetheless, we think that there is a potential of avoiding to select which minutia to empower with SAFE features, because when a minutia is unique alone (without constellation information), this is indicated by the fraction of the energy in the non-DC components of SAFE features as compared to the total energy. The strength of the object property of a minutia neighborhood changes gradually as the distance to a core and/or delta changes. The rate of change is also individual to the finger.

We have introduced a feature vector which describes the object properties of an image neighbourhood. We have justified the object property concept in that if SAFE features (non-DC components) are strong it means that there is a (mathematical) singularity of the image neighborhood, whose location is
4.3. DISCUSSION

intrinsically and uniquely defined by the image content in a translation and rotation invariant manner. The non-DC components are silent in case the image neighborhood is part of a texture. This gives a higher precision as to which property of the image neighborhood a SAFE feature vector gives a strong identity and which unique location it intrinsically defines. We argue that this precision is in contrast to popular descriptors, e.g. SIFT [51], SURF [52], LBP [53], which suggest being universal, but yet have not been studied sufficiently w.r.t. whether or not they can differentiate between a texture neighborhood which has no intrinsic and unique location defined by itself and an object neighborhood which has such an intrinsic and unique point.

We think that the differentiation of SAFE features between the wave and the particle nature of a neighborhood can be useful to other applications too. We have presented experimental evidence supporting this view in the form of identity recognition by periorcular regions of faces. We have suggested that the particle nature description capacity of SAFE features explains why these features have improved the performance of Gabor filter magnitude responses, which are well known descriptors for being powerful in quantifying texture (wave) neighborhoods. Nonetheless, further studies should be conducted to predict which applications can benefit significantly from them without extensive experiments. An alternative is to always use them in complement to texture features, if such knowledge is not acquired.

After the publication of conference papers [C1], [C2], we have noticed a typo in the rotation compensation formula. The correct version of it is available in the journal submission version, paper [C] eq. (18).

For pedagogical reasons, we have also made (non-essential) changes in the presentation of the frequency information in Paper [C] in comparison to the (published) conference Papers ([C1], [C2], which are comprised in Paper [C]). These were first, the removal of the outlier frequency estimations, based on a priori knowledge that ridge periods in 500 dpi fingerprints are between 6 and 13 pixels. Second, the color has been remapped such that increasing spatial frequency corresponds to increasing electromagnetic wave frequencies of color. Previously increased ridge periods were mapped to increasing wavelengths of color. The pedagogy is in the mnemonics, attempting to support the human memory to remember the legend, even without an explicit color chart. A knowledge of how color progresses in rainbow \(^1\) would then suffice to remember the essentials of our frequency (as well orientation) legends. The two hue ranges representing the same dense frequency map of a fingerprint are shown in Fig. 4.6 including the respective color legends.

\(^1\)...or the chromatic diagram standardized by the Committee International de l’Eclairage
Figure 4.6: Top: The same frequency estimation (of a fingerprint) but represented in two different color representations. Bottom: The color legends representing the spatial frequencies when these are mapped to the two hue ranges used above. The period changes linearly between [2.5, 15] pixels.
Chapter 5

Final Conclusion

Understanding Image Processing by experts of Biometrics and Forensics is increasingly important for a more effective analysis of the image data. This is because special tools can be designed for specific problems that were previously not possible to obtain automatically, but now possible due to progress in image analysis. Not the least it is a legitimate reason to avoid to use methods from other applications of image processing without taking into account the needs of the fields of biometrics and forensics. In fact, some of these problems are fundamental to image processing, a contribution to which would make a difference in other applications of image processing too. Accordingly, developing fundamental image analysis methods using the field as test bed should be fruitful even to image processing science.

We have focused on automatic feature extraction when key points are known with special emphasis to quantitate local patterns by (visual) pattern models imaginable (which curve family), interpretable precisely (which member of a family and at what confidence/quality), and verifiable (humans can see identify the specific symmetric curves themselves in a fingerprint) by humans. This allowed to measure the quality of the estimated model parameters too, based on deviations from models of symmetric curve families. We have tested this model in fingermark recognition, and face recognition by periocular regions.

A central question in this study was thus whether or not continuous orientation maps around minutiae have significant information not captured by minutia plus minutia orientation, i.e. can this information be used to improve the individual recognition power of a minutia. If so, evidently this must be extracted automatically and offered to human experts and machines to improve the quality of the forensic decision (identification or not). We have shown that indeed this neighborhood data carries significant identity information. We have shown how the orientation information around minutiae can be rep-
resented compactly by finite expansion in terms of predefined symmetric curve families, with promising identification performance.

Periocular images were used for testing the persistence of the usefulness of SAFE features across different image qualities and scanning resolutions. In fingermark application the image resolution was constant (500 dpi) but the noise was omnipotent and the size of the useful fingermark sizes were small. In periocular region recognition, our experiments gave also support to the existence of information with object/particle nature beside texture/wave nature in such images, showing complementarity to texture descriptors.

The features are shown to be rotation invariant in the mathematical sense, i.e. if the rotation difference is known, there is no need to rotate the image and recalculate the features from the scratch. The new features can be obtained by complex multiplications of the original features. Additionally, the log-spiral features are absolutely invariant to rotations, i.e. there is no change in these features at all.

Because they can be imagined, interpreted, and verified by humans it can be postulated that these features encode what humans see in neighborhoods of key points, regardless if the human visual system actually uses this encoding internally. The encoding offers at least a practicable way of transforming human visual perception to machines, which they too can use to pull out individual fingermarks from databases of fingerprints.

We would like to emphasize the need for incorporating the feature vector containing the extended features into the fingerprint matching algorithms as well. The current work puts forward contributions in automatic image processing as an assistance for forensic examiner but to reach its full potential even novel matchers should be developed to use this information when matching constellations of minutiae.

Automatically extracted orientation and frequency maps, ridge counting, rotation invariant features were evaluated on publicly available data. This opens possibility for comparison of future image based features.

Although we have sketched a possible scenario on how these features can be incorporated to forensic "calculators" for semi-automatic matching, software efficiently supporting these, including GUIs, are yet to be developed.

To evaluate the performance of feature descriptors, we have provided ground truths for the SD27 database, making the annotation as deep as minutiae level publicly available. The dataset was lacking the ground truth at key points level (especially minutiae). We believe that this will facilitate in depth studies of new features, and evaluating the identification power of them, separate from the constellations in which the key points are present.

The proposed features can be upgraded with scale invariance augmenting the application area of it. As of yet SAFE features are "only" rotation and translation invariant, given the key points. The presumption is realistic in fingerprints recognition, because human experts are currently extracting the key points in fingermarks, and machines do that in tenprints. Two prior studies
have shown that cores and deltas can be detected in tenprints, by using features corresponding to parabolic curves and triangular-hyperbolic curves of SAFE features in tori peripheral to a key point [49], and minutiae can be detected in tenprints by ditto, but by parabolic curve families in tori close to the key point, [46]. This raises the hope that SAFE features can be also used in a way that is it can grade which points are cores, deltas and minutiae in fingermarks and offer those that are above a certain quality, in the future.

Our preliminary studies (not published) show that matching forensic shoeprints with object-based features show considerable room for improvement, suggesting that SAFE features should possibly be complemented by texture features, e.g. Gabor features, [54], [55] as the results on periocular region recognition indicates Paper [C].

Also, image processing improvements could include local orientation and frequency map (and even image) enhancements, the delimitation of which can be automatic or suggested by an expert. In our study we have focused on providing the enhancements for the entire fingermark.

Aggregating feature vectors of single points results to a full fledged identification system (at finger/person level), beyond being a forensic "calculator" requires more study. In case of periocular recognition we were able to join the novel features to existing features on a grid of points. However, the grid point distances were regular, rigid, with correspondence between the two tested grids available, and the used grids were provided to have the same rotation between the compared periocular data (alignment of eyes to horizontal direction is feasible), whereas the same properties of fingermarks/tenprints were the opposite, increasing the difficulty of a similar development of a matcher for fingermarks with resources available to our project.

One way of incorporating SAFE features can be envisaged to be done by using Commercial, Off The Shelf (COTS) software to match minutia constellations one component of SAFE vectors at a time. Since the suggested features represent angles which also exist in a minutia constellation a COTS may be "convinced" to perform a matching by replacing the angles of minutiae with those of SAFE features. However, this is not as straightforward for all SAFE features. Only parabolic symmetries would allow themselves for such a replacement, because rotating the finger globally induces the same rotation in them, whereas the other features must somehow inverse compensated for the N-folded symmetries they inherently have. By way of example, if a rotation of images with $\varphi t$ is assumed between a tenprint/fingermark, the angles corresponding to triangular/hyperbolic curve components of SAFE features, $n = 1$, must be translated with $3\varphi t$, see eq. (18) of Paper [C]. A COTS will however internally translate the tenprint angles with $\varphi t$ towards those of the fingermark (instead of with $3\varphi t$) before matching them. Furthermore, the spiral family angles, $n = -2$ should not be translated internally by the COT, but being a black-box, these angles too will be rotated (with $\varphi t$), if no additional measures are present. How the data should be preprocessed such that they can
be blindly used by COTS should therefore be the subject of a future study, because it would allow to easily use existing infrastructure of software used by forensic experts.

Our enhanced orientation maps can also be considered as potential contributors to extraction of tenprint sub-regions, similar to [56]. In the latter, orientation fields were artificially constructed by interpolation using (only) minutiae directions which are very sparse. This procedure was used to extract the most probable box of a minutiae subset (rather than using all minutiae) in the tenprint, which is subsequently shown to improve fingerprint matching. The improvement effect of such a procedure could be studied by using real (instead of interpolated) orientation fields, e.g. the enhanced orientation fields of fingerprints suggested in Paper [B], possibly corrected by an examiner.
Bibliography


BIBLIOGRAPHY


