Error propagation in pattern recognition systems: Impact of quality on fingerprint categorization

Jakub Mocny
University of New Hampshire, Durham

Follow this and additional works at: https://scholars.unh.edu/dissertation

Recommended Citation
https://scholars.unh.edu/dissertation/412

This Dissertation is brought to you for free and open access by the Student Scholarship at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact nicole.hentz@unh.edu.
ERROR PROPAGATION IN PATTERN RECOGNITION SYSTEMS: IMPACT OF QUALITY ON FINGERPRINT CATEGORIZATION

BY

JAKUB MOCNY

BS, Technical University of Gdansk, 2000
MS, Technical University of Gdansk, 2000
MS, University of New Hampshire, 2004

DISSERTATION

Submitted to the University of New Hampshire in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

In

Engineering: Systems Design

December, 2007
This thesis has been examined and approved.

Dissertation Director, Michael J. Carter, Electrical and Computer Engineering

Filson H. Glanz, Electrical and Computer Engineering

Ernst Linder, Mathematics and Statistics

Richard A. Messner, Electrical and Computer Engineering

W. Thomas Miller, III, Electrical and Computer Engineering

11-27-2007

Date
ACKNOWLEDGEMENTS

This work has been partially funded by the Graduate School Dissertation Fellowship Award.

Above all I would like to stress that the tremendous task I embarked on a few years ago would not have come to a successful completion without help and support from the faculty and staff of the Electrical and Computer Engineering and Systems Engineering Programs. I would like to express my sincere gratitude to Professor John LaCourse, Chairman of the Electrical and Computer Engineering Department and to Dr. Joseph Klewicki, Dean of the College of Engineering and Physical Sciences as well as to Dr. Harry Richards, Dean of the Graduate School, and to Dr. Cari Moorhead, Associate Dean of the Graduate School, for their financial support in the Academic Year of 2005/2006.

I am particularly indebted to my Dissertation Advisor, Professor Michael Carter for the countless hours of fruitful discussions. His support and guidance were invaluable. I would also like to extend special thanks to all the members of the Committee for the time and effort they spent reviewing the dissertation and for the insightful discussions that helped me find different perspective on the problems at hand.

The warmest thanks go to my wife Ania for her understanding, patience and moral support throughout all these years, to my parents and aunt Ula for the words of encouragement when I needed them most. And last but not least I would like to thank the Boy family, especially the Colonel, for making it possible for my wife and me to feel at home.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii

LIST OF TABLES vi

LIST OF FIGURES vii

LIST OF ACRONYMS xii

ABSTRACT xiii

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Why is Fingerprint Quality Important?</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Outline of Chapters</td>
<td>2</td>
</tr>
<tr>
<td>2. PROBLEM FORMULATION</td>
<td>4</td>
</tr>
<tr>
<td>2.1 Pattern Classification Paradigm</td>
<td>4</td>
</tr>
<tr>
<td>2.2 Quality Aware Classification</td>
<td>8</td>
</tr>
<tr>
<td>3. QUALITY OF DIGITAL FINGERPRINT IMAGES</td>
<td>14</td>
</tr>
<tr>
<td>3.1 Salient Features</td>
<td>14</td>
</tr>
<tr>
<td>3.2 Defects - Qualitative Approach</td>
<td>18</td>
</tr>
<tr>
<td>3.3 Models of Quality Defects</td>
<td>24</td>
</tr>
<tr>
<td>3.4 Defects - Quantitative Approach</td>
<td>32</td>
</tr>
<tr>
<td>4. FINGERPRINT RECOGNITION SYSTEM BUILDING BLOCKS</td>
<td>40</td>
</tr>
<tr>
<td>4.1 Pattern Collection</td>
<td>40</td>
</tr>
<tr>
<td>4.2 Preprocessing</td>
<td>46</td>
</tr>
<tr>
<td>4.2.1 Segmentation</td>
<td>47</td>
</tr>
<tr>
<td>4.2.2 Enhancement</td>
<td>62</td>
</tr>
<tr>
<td>4.3 Feature Extraction</td>
<td>65</td>
</tr>
<tr>
<td>4.4 Pattern Classification</td>
<td>80</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. CERTAINTY OF ORIENTATION ESTIMATE</td>
<td>86</td>
</tr>
<tr>
<td>5.1 Orientation Estimation Methods</td>
<td>86</td>
</tr>
<tr>
<td>5.2 Orientation Estimation in Presence of Defects</td>
<td>92</td>
</tr>
<tr>
<td>5.3 Orientation Uncertainty – Quantitative Approach</td>
<td>101</td>
</tr>
<tr>
<td>6. ANALYSIS OF THE EFFECT OF QUALITY DEFECTS ON ORIENTATION ESTIMATE</td>
<td>111</td>
</tr>
<tr>
<td>6.1 Principle of Ridge Orientation Estimation</td>
<td>111</td>
</tr>
<tr>
<td>6.2 Estimation in the Presence of Noise</td>
<td>117</td>
</tr>
<tr>
<td>6.3 Estimation in the Presence of Anisotropic Defects</td>
<td>125</td>
</tr>
<tr>
<td>6.4 Finite Window and Sampling Effects</td>
<td>135</td>
</tr>
<tr>
<td>7. CONCLUSIONS AND FUTURE WORK</td>
<td>137</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>141</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>148</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 3.1. The most common minutiae [Maltoni et al., 2003]. 17
Table 3.2. Fingerprint defects. 24
Table 4.1. Parameters of fingerprint sensors. 43
Table 4.2. The live-scan technology overview. 45
Table 4.3. Local maxima of $\frac{\theta_{d,\lambda}}{c_\lambda}$. 71
Table 4.4. Singular point-based categorization. 82
Table 4.5 Performance in 3- and 4-class categorization task. 85
Table 5.1. Estimated local ridge orientations for the patterns shown in Figure 5.2 - Figure 5.5 and for 5 different estimation methods. 98
Table 5.2. Comparison of the coherence of normalized $c_2$ and non-normalized gradients $c_1$ for the cases presented in Figure 5.7. 102

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
LIST OF FIGURES

Figure 2.1. Pattern recognition system – block diagram. 5
Figure 2.2. The marginal probability \( p(x) \) – a), and the decision boundary – b). 6
Figure 2.3. The block diagram showing the pattern recognition system in which quality information is used as side information. 10
Figure 2.4. The sample \( I \) contains the quality information that can be included in classification. The class \( \omega_i \) generates features \( x \) that are embedded in the raw pattern \( I \) possibly distorted as a result of low quality. 10
Figure 2.5. The marginal probability \( p(x) \) - a), and \( p(x|I) \) - b). 12
Figure 3.1. The most common fingerprint classes NIST DB4; arch a), tented arch b), left loop c), right loop d), whorl e). 15
Figure 3.2. Core and delta singular points NIST DB4; tented arch a) with core and delta SPs, left loop b), right loop c) and whorl d) with two core SPs (○) at the center and two delta SPs (▲) on the periphery of the impression. 16
Figure 3.3. Images from FVC2000 DB1b (a,b) and FVC2002 DB2b (c,d) - defect 1. 20
Figure 3.4. Synthetic images from FVC2000 DB4b (a,b) and live scans FVC2004 DB1b (c,d) - defect 2. 20
Figure 3.5. Images from FVC2000 DB2b (a,b) and FVC2004 DB1b (c,d) - defect 3. 21
Figure 3.6. Images from FVC2000 DB1b (a,b) and FVC2000 DB3b (c,d) - defect 4. 21
Figure 3.7. Images from FVC2004 DB3b (a,b), and NIST special database 4 (c) - defect 5. 22
Figure 3.8. Images from FVC2000 DB1b (a,b), DB2b (c), DB2b (d) - defect 6. 22
Figure 3.9. Images from FVC2000 DB2b (a), DB3b (b), FVC2002 Db3b (c) - defect 7. 23
Figure 3.10. Images from FVC2000 DB1b (a), FVC2002 DB3b (b), FVC2004 DB1b (c) - defect 8. 23
Figure 3.11. Sample images; underlying pattern is a sine wave and the defects are modeled by a multiplicative mask composed of a superposition of Gaussian functions. 27
Figure 3.12. Anisotropic defect a) consisting of a superposition of circularly symmetric Gaussian functions and image \( s_{\text{inv}}(x,y) \) after applying a multiplicative mask \( m_{\text{inv}}(x,y) \) b). 28
Figure 3.13. Synthetic images illustrating the effect of the contrast compression; the
original image a), low contrast, high mean intensity, $c_{cr}=0.3$, $d_{cr}=0.7$ b), low
contrast, low mean intensity, $c_{cr}=0.3$, $d_{cr}=0$ c).

Figure 3.14. Synthetic images; original $s_{m}(x,y)$, $\alpha_{pr}=1$ a), low pressure $\alpha_{pr}=0.3$ b), high
pressure $\alpha_{pr}=3$ c).

Figure 3.15. The plot of $\sin(2\pi f_{x}x)$; the widths of the ridges can be found by measuring
the length of the line segment from the intersection of the horizontal line with
the positive slope to the first following negative slope, according to (3 13); $n$
denotes a sample number.

Figure 3.16. Maximum of local gradient of the $f(x,y)$ in a (log-log) plot of scale parameter
$\sigma$ a) and $\|\mathbf{g}_{\text{max}}\|$ where the maximum of the gradient is located in $r=\sigma$ from the
center b).

Figure 3.17. Synthetic image showing a sine wave distorted according to (3-23) a), and
an amplitude spectrum without DC component b).

Figure 3.18. The effect of different values of $\alpha_{pr}$ on Fourier spectrum; wider ridges left
and narrower ridges right correspond to different degrees of concentration of
power spectrum around 0.

Figure 4.1. Image from NIST-4 database a), and a magnified fragment b).

Figure 4.2. Horizontal profile of intensities a), taken at the mid-section of the image in
Figure 4.1, and b) the corresponding response of the mean, standard
deviation, $d_{1}'$ and $d_{2}'$.

Figure 4.3. Magnified fragment of the fingerprint from Figure 4.1 a) and histogram of
intensities b).

Figure 4.4. Magnified fragment of the fingerprint from Figure 4.1 a) and histogram of
intensities b).

Figure 4.5. Histograms of measures calculated for the image shown in Figure 4.1.

Figure 4.6. Segmentation measures, where low intensity denotes low value and high
intensity high value for a) mean, b) standard deviation, c) and d) .

Figure 4.7. Segmentation results obtained when using fixed threshold: a) mean (m<0.44),
b) standard deviation ($\sigma>0.5$), c) $d_{1}'<0.5$, d) $d_{2}'<0.5$. Image f05 from NIST 4.
Figure 4.8. Likelihood estimates obtained using Gaussian Parzen windows for calculated for image from Figure 4.1 on a rectangular grid of $A$, $c=9$, and $c_\ell=19$ pixels (5 pixel overlap); a) $p(c_\ell \mid F)$ and b) $p(c_\ell \mid \overline{F})$. Compare with histogram c) from Figure 4.5.

Figure 4.9. MAP segmentation based on a,c) standard deviation, b,d) based on d-prime 1. (a,b - image f05 from NIST 4, and c,d - image 110_7 from FVC2000 DB3b).

Figure 4.10. Scatter diagram showing relationship between points in the cliques illustrating the different statistical properties for ridge pattern a), and noisy background b).

Figure 4.11. Examples of images, a) FVC 2000 DB 3 a, and b) FVC 2000 DB 3 b.

Figure 4.12. Comparison of the outcomes of segmentation using d-prime 1 a,c) and 2-cliques b,d).

Figure 4.13. The results of image enhancement using the technique presented in [Maltoni et. al. 2003] (the second column) and the modified approach in which a bandpass filtered signal was used to construct an enhancement filter (last column). First row – ridges with breaks, second row – a good quality image, next two rows – subimages taken from a low quality fingerprint image and the bottom row – a synthetic image.

Figure 4.14. Butterworth bandpass filter of the 9th order.

Figure 4.15. Fingerprint images (FVC 2004 Database 1 b); different impressions of the same fingerprint ordered in decreasing perceptual quality.

Figure 4.16. Cost function $\theta_0$ as a function of block size.

Figure 4.17. Histograms of block sizes obtained using stage 1 without overlap between blocks (o – FVC2000 DB3b, x – FVC2004 DB1b).

Figure 4.18. Smooth orientation field estimation; first stage a) – block size is 17, no overlap, $\theta_0=9.26$, after second stage b) – block size is 33, grid is 17 (overlap=8), $\theta_0=6.03$.

Figure 4.19. Smooth orientation estimate; a) coherence of gradient vectors $c_k$, b) orientation field $\theta_k$, c) local orientation difference $\theta_{d,k}$, and d) $\theta_{d,k}/c_k$ (for image in Figure 4.20).

Figure 4.20. Local maxima of the ratio of $\theta_{d,k}/c_k$ a) and singular point detected using Poincaré index based method [Bazen and Gerez, 2002] b).
Figure 4.21. Smooth estimate of orientation field using MRP \( \theta_b = 4.73 \) (block size is 17 pixels with no overlap).

Figure 4.22. Examples showing singular points detected using Poincaré index-based method [Bazen and Gerez, 2002]; from left to right FVC2000 DB3b, FVC2004 DB1b, NIST Database-4.

Figure 4.23. Fingerprint examples with superimposed OFFC; tented arch a), whorl b).

Figure 4.24. Examples of \( \cos(y_i) \) - a) and OFFC - b) for an arch pattern.

Figure 4.25. Examples of \( \cos(y_i) \) - a) and OFFC - b) for a left loop pattern.

Figure 4.26. Examples of \( \cos(y_i) \) - a) and OFFC - b) for a whorl pattern.

Figure 5.1. Results of orientation field estimation using gradient method; a) block size is 17 (no overlap), b) block size is \( c_b = 33 \), grid is \( c = 17 \) (overlap of 8 points).

Figure 5.2. Orientation signatures for a good quality pattern (impression 103_1 from FVC 2004 DB3b).

Figure 5.3. Orientation signatures for a low quality pattern (impression 105_6 from FVC 2000 DB3b).

Figure 5.4. Orientation signatures for a sample containing over-emphasized ridge pattern (impression 102_8 from FVC 2004 DB3b).

Figure 5.5. Orientation signatures for a ridge pattern containing directional defect (impression 108_1 from FVC 2000 DB3b).

Figure 5.6. Boundary of background and fingerprint ridge pattern; coherence of gradients is high.

Figure 5.7. Figure shows scatter diagrams of gradients: for the entire image (Figure 5.6) a), bottom right quadrant with the ridge pattern b), and the top left quadrant containing background noise c). The diagrams d), e), and f) depict corresponding gradient orientation histograms.

Figure 5.8. Orientation signature obtained from the response of Gabor filters a) and the corresponding quality descriptors \( f(s) = 0.029 \) and \( f'(s) = 0.353 \). The corresponding fragment of a fingerprint image b).

Figure 5.9. Signal \( x(n) \) a), and variance of frequency localization b).

Figure 5.10. Variance of orientation estimates; \( (x) \) – sample variance, \( (\hat{\theta}) \) – predicted using (5), and \( (\Theta) \) function of coherence of gradients.

Figure 5.11. Test image distorted by additive Gaussian noise (with zero mean); with variance equal to 0.001 a), 0.01 b) and 0.1 c).
Figure 6.1. MSE fit of the line Ax+By+C=0 to the data.

Figure 6.2. The orientation signature for the three cases of defect orientation and shape;

- $\phi=0, \sigma_x=1, \sigma_y=7$ (first row), $\phi=-10, \sigma_x=1, \sigma_y=7$ (second row) and $\phi=-10, \sigma_x=5, \sigma_y=7$ (the last row).

Figure A.1. Comparison; scatter diagram of standard deviation of intensities and $c_1$ a), and $c_2$ b).

Figure A.2. Scatter plot $c_1$ vs. $c_2$.  

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-D</td>
<td>two dimensional</td>
</tr>
<tr>
<td>DB</td>
<td>database</td>
</tr>
<tr>
<td>DC</td>
<td>direct current, average</td>
</tr>
<tr>
<td>DFT</td>
<td>Discrete Fourier Transform</td>
</tr>
<tr>
<td>FBI</td>
<td>Federal Bureau of Investigation</td>
</tr>
<tr>
<td>FFT</td>
<td>Fast Fourier Transform</td>
</tr>
<tr>
<td>FT</td>
<td>Fourier Transform</td>
</tr>
<tr>
<td>FVC</td>
<td>Fingerprint Verification Competition</td>
</tr>
<tr>
<td>IAFIS</td>
<td>Integrated Automated Fingerprint Identification System</td>
</tr>
<tr>
<td>ICA</td>
<td>independent component analysis</td>
</tr>
<tr>
<td>IID</td>
<td>independent identically distributed</td>
</tr>
<tr>
<td>IQM</td>
<td>Image Quality Measure</td>
</tr>
<tr>
<td>MAP</td>
<td>maximum $a$ posteriori</td>
</tr>
<tr>
<td>MMSE</td>
<td>minimum mean squared error</td>
</tr>
<tr>
<td>MSE</td>
<td>mean squared error</td>
</tr>
<tr>
<td>NIST</td>
<td>National Institute of Standards and Technology</td>
</tr>
<tr>
<td>OFFC</td>
<td>orientation field flow curve</td>
</tr>
<tr>
<td>PCA</td>
<td>principal component analysis</td>
</tr>
<tr>
<td>PPFT</td>
<td>Pseudo-Polar Fourier Transform</td>
</tr>
<tr>
<td>SNR</td>
<td>signal to noise ratio</td>
</tr>
<tr>
<td>SP</td>
<td>singular point</td>
</tr>
</tbody>
</table>
ABSTRACT

ERROR PROPAGATION IN PATTERN RECOGNITION SYSTEMS:
IMPACT OF QUALITY ON FINGERPRINT CATEGORIZATION

by

Jakub Mocny

University of New Hampshire, December, 2007

The aspect of quality in pattern classification has recently been explored in the context of
biometric identification and authentication systems. The results presented in the literature indicate
that incorporating information about quality of the input pattern leads to improved classification
performance. The quality itself, however, can be defined in a number of ways, and its role in the
various stages of pattern classification is often ambiguous or ad hoc.

In this dissertation a more systematic approach to the incorporation of localized quality
metrics into the pattern recognition process is developed for the specific task of fingerprint
categorization. Quality is defined not as an intrinsic property of the image, but rather in terms of a
set of defects introduced to it. A number of fingerprint images have been examined and the
important quality defects have been identified and modeled in a mathematically tractable way.
The models are flexible and can be used to generate synthetic images that can facilitate algorithm
development and large scale, less time consuming performance testing. The effect of quality
defects on various stages of the fingerprint recognition process are examined both analytically
and empirically. For these defect models, it is shown that the uncertainty of parameter estimates,
i.e. extracted fingerprint features, is the key quantity that can be calculated and propagated
forward through the stages of the fingerprint classification process. Modified image processing
techniques that explicitly utilize local quality metrics in the extraction of features useful in
fingerprint classification, such as ridge orientation flow field, are presented and their performance
is investigated.
CHAPTER 1

INTRODUCTION

1.1 Why Is Fingerprint Quality Important?

Image quality plays an important role in the assessment of performance of imaging devices and image processing algorithms. Often it is used in the context of the human visual system, in which case the quality metric is designed to mimic the perception of a human observer [Avcibas et al., 2002]. The computational measures can be mapped to correlate well with the human-based model. If the original image is used as a reference for computing the quality, then such a metric is referred to as full-reference [Shnayderman, 2006]. A more difficult scenario arises when only one image is available and we are interested in a way of assessing its quality. In this, so-called no-reference approach [Shnayderman, 2006], it is necessary to examine properties of the image. It can be accomplished for example, based on analysis of the Fourier spectrum, as in the case of the IQM algorithm [Nill and Bouzas, 1992] designed originally for aerial images and subsequently used in the case of fingerprints. It is required that certain invariance properties apply, i.e. we are interested in determining the quality irrespective of the image content. However, this can be achieved to a limited extent. The aerial images and fingerprint images have different power spectra and require different approaches. It follows that it makes sense to compare quality of images of the same type only. In this thesis an attempt is made to discover how the quality influences the performance of automated recognition systems. While such influence is generally acknowledged, its mechanism is not known precisely. The approaches to quality evaluation, mentioned above, can be designed in such a way that they predict the classification outcome, as in [Tabassi et al., 2004], [Chen et al., 2005] and [Chen et al., 2006]. The predictive properties vary between systems. This is intuitively justified since quality appears to be highly subjective. Quality necessarily has to depend on what we are looking for in the scene and what our methods of finding it are. Then the quality metric can act as a measure of our ability of finding interesting details in an analyzed scene. This requires quality to be constructed differently for different image processing and classification algorithms. In a typical implementation in pattern recognition systems, features are extracted from the image and are used in a classification stage to determine
the class of the pattern present in the image. In fingerprint classification the question can concern the identity of the owner of the fingerprint or the global characteristics of the ridge-furrow pattern. Typically, the features used in classification are composed of lower level features. The case study used in this thesis relies on local ridge orientations, which constitute basic components of the orientation field that has characteristics useful for categorization task. The quality impairments, referred to in this thesis also as quality defects, can be identified and represented using models, as shown in Chapter 3. The effect of different quality defects can be quantified, and this is shown for low level feature extraction, cf. ridge orientation estimation methods. Following the premise of context-dependency of quality, alternative quality metrics are proposed in Chapter 5 for representative orientation estimation methods. The quality is related to the uncertainty of the orientation estimate. Higher level features in some way are influenced by quality as well. Smooth orientation field estimation incorporates local uncertainty and the assumption of smooth changes in orientation. However, this does not answer the question of what the uncertainty of such an estimate is. This question remains yet to be addressed. From examples shown in Chapter 2, it appears that the pattern quality expressed as an uncertainty of the feature estimates can serve not only as a predictor of the classification outcome as proposed in the examples in the literature, but it can be used explicitly in classification. It appears plausible in any conceivable classification scenario, given that it can be represented in the form shown in Section 2.2.

1.2 Outline of Chapters

In Chapter 2, based on a Bayesian framework, it is justified why using the quality information is important in classification. An example, in which a hypothetical feature vector is used, is presented. Two different classification outcomes are demonstrated. In the first one, the feature vector is assumed to have a value associated with the maximum of the feature signature, expressed as a conditional probability of a feature given the raw pattern \( I \), \( p(x|I) \). In the second one, the knowledge of the features is assumed to be uncertain in which case the entire \( p(x|I) \) is used.

Chapter 3 introduces the reader to the properties of the fingerprint patterns that are useful in classification. Subsequently, examples of quality-impaired images are presented, from which a defect model is derived and presented. Selected properties of the modeled defects are presented in reference to the phenomena that inspired them.
The building blocks of a categorization algorithm are presented in Chapter 4. They are subsequently used in evaluation of the impact of quality on the classification task. Certain innovations are introduced to the methods proposed in the literature, and new methods are presented.

Chapter 5 discusses discrete implementations of the low level feature extraction, i.e. local ridge line orientation. Fourier Transform-based methods can be easily modified to include local frequency estimation. Local quality descriptors are proposed for different methods for orientation estimation. With respect to the approaches shown in the literature, it is herein proposed to explicitly define the quality based on the output of a given processing method rather than trying to obtain an objective quality metric. It is achieved through the measures based on properties of the orientation signature.

Chapter 6 contains useful insights into the orientation estimation methods. In the presence of selected defects emulated using models from Chapter 3, the performance of such estimators is analyzed and described using closed-form expressions. The Fourier Transform (FT) is used as a tool for the analysis. It allows stating the common principle underlying the orientation estimation methods.

Chapter 7 contains the summary of the thesis findings, unsolved problems, and suggestions for the future work.
CHAPTER 2

PROBLEM FORMULATION

In this chapter a theoretical framework underlying pattern classification is presented. Section 2.1 introduces the reader to the Bayesian approach in which the features are used to determine the \textit{a posteriori} probability that a pattern belongs to a given class. It is followed in Section 2.2 by a formulation of the pattern classification problem in a form that allows incorporating information about quality into the decision process. The enclosed examples emphasize the difference between the two approaches.

2.1 Pattern Classification Paradigm

A pattern recognition system can be described as a sequence of operations in which a sample containing a pattern of interest is first collected. The pattern of interest is typically embedded in a background of patterns that from the standpoint of pattern classification can be considered as noise. One of the first preprocessing operations therefore is separation of the pattern of interest from background. In the image processing domain, this process is referred to as segmentation. For example, in the fingerprint recognition system the fingerprint silhouette is separated from background by taking advantage of distinct properties that characterize a fingerprint. Further preprocessing techniques may involve image enhancement; typically the raw pattern is modified to facilitate feature extraction. In the case of fingerprints, examples of preprocessing include spatial bandpass filtering, image binarization, and ridge thinning. Bandpass filtering improves the visibility of ridge pattern, the noise is reduced and breaks in ridges can be eliminated. The resulting image is much more suitable for binarization followed by ridge thinning. Ridge thinning, also referred to as skeletonization, facilitates minutiae extraction. Fingerprint minutiae are probably the most important features used in fingerprint recognition.

In any pattern recognition problem, and especially in the case of biometric systems, it is crucial that the set of features is unique, stable over time, and that it can be reliably extracted from a sample of reasonable quality. The types of features are dictated by the type of classification task. For instance, a categorization task requires features that provide discriminatory information to classify an object into its general categories of planes, cars, trees,
etc. These categories can be further subdivided into more specific classes such as a type of a plane, a car, and species of a tree, within which, in turn, one can perform recognition of an individual object.

In the case of fingerprints, a number of categories can be identified based on global characteristics of ridge flow. Useful features for this task consist of orientation field, orientation field flow curves (OFFCs) and singular points (SPs). On the other hand, in order to determine the identity of a person one needs a more specific approach in which individual properties of each fingerprint are examined. This classification task requires more information such as minutiae, SPs, ridge shape, and sweat pores. Their locations and types are essential for the comparison of fingerprints. In the final stage of the pattern recognition the pattern is assigned to one of the classes. Often a matching score is used to denote the strength of a match. In practice, a number of different approaches to classification can be encountered in the literature ranging from syntactic, rule-based to probabilistic.

![Pattern recognition system - block diagram](image)

Figure 2.1. Pattern recognition system - block diagram

The framework for pattern classification can be formulated using Bayes theorem (2-1).

\[
p(\omega_i \mid x) = \frac{p(x \mid \omega_i) p(\omega_i)}{p(x)} = \frac{p(x \mid \omega_i) p(\omega_i)}{\sum_j p(x \mid \omega_j) p(\omega_j)}
\]

(2-1)

It follows that in order to accomplish a classification task in an optimal fashion, it is necessary to know the underlying probabilities. Based on this information, a classifier function that divides a feature space into subsets corresponding to different classes can be found. The different approaches to classification can be viewed as different approaches to design of the classifier function. The latter does not have to be explicit. In the context of pattern classification, Bayes theorem expresses a probability that a given measurement - a feature \( x \in X \), where \( X \) is a set of features, belongs to a class \( \omega_i \) from some set of classes \( \Omega \) suitable for a given problem.
Often \( X \subseteq R^d \); i.e., the feature set is a subset of the \( d \)-dimensional space of real numbers, and \( x \) is then a \( d \)-dimensional real vector in this space.

The illustration below presents a solution to a 2-class classification problem. This example is then used in the following section where it is modified to account for the uncertainty associated with the features. Given the probabilities, an optimum decision boundary is found. The features are 2-dimensional (2-D) vectors. The likelihoods \( p(x|\omega) \) are a Gaussian function for \( \omega_1 \) and a mixture of Gaussians for \( \omega_2 \). The prior class probabilities are assumed to be equal.

\[
\begin{align*}
p(x|\omega) &= \mathcal{N}(\mu, \Sigma) \\
\mu &= [20 \ 10]^T, \quad \Sigma = \begin{bmatrix} 361 & 0 \\ 0 & 361 \end{bmatrix}
\end{align*}
\]

and \( p(x|\omega_2) \) consists of two clusters and is defined as a mixture of Gaussians:

\[
p(x|\omega_2) = 0.5 \cdot \mathcal{N}(\mu_{21}, \Sigma_{21}) + 0.5 \cdot \mathcal{N}(\mu_{22}, \Sigma_{22}) ,
\]

Figure 2.2. The marginal probability \( p(x) - a \), and the decision boundary - b).

The marginal probability of \( x \), \( p(x) \) is shown in Figure 2.2 a). The class conditional probability \( p(x|\omega) \) corresponds to a single maximum or a single cluster and is defined as a bivariate Gaussian,

\[
p(x|\omega) = N(\mu, \Sigma)
\]

and \( p(x|\omega_2) \) consists of two clusters and is defined as a mixture of Gaussians:
\[
\begin{bmatrix}
\mu_{21} &=& \begin{bmatrix} -3 & 3 \end{bmatrix}^T \\
\mu_{22} &=& \begin{bmatrix} -3 & -3 \end{bmatrix}^T \\
\Sigma_{21} &=& \Sigma_{22} = \begin{bmatrix} 225 & 0 \\ 0 & 225 \end{bmatrix}.
\end{bmatrix}
\]

The decision boundary is shown in Figure 2.2 b) and represents a line of equal probability of misclassification to \(\omega_1\) and \(\omega_2\) given the MAP decision rule (2-4).

One way of looking at the pattern recognition system is to consider it to be a sequence of transformations between spaces of different dimensionality. A pattern collection can be viewed as a mapping, denoted by \(S\), representing a measurement a physical quantity, such as a voltage \(U\), resulting from the presence of an electric charge between the finger and the surface of a capacitive sensor:

\[
I = S(U).
\]  \hspace{1cm} (2-2)

In our work \(S\) represents a fingerprint sensor. Only features that are important for classification are retained from the raw pattern \(I\). This is useful in order to limit the dimensionality of a problem to a more manageable size. In the equation (2-3) below, the operators include sample collection \(S\), preprocessing \(P\) and feature extraction \(F\).

\[
x = F \circ P(I) = F \circ P(S(U))
\]  \hspace{1cm} (2-3)

The feature \(x\) is then classified according to the rule (2-4):

\[
\text{choose } \omega_i \text{ if } p(\omega_i \mid x) > p(\omega_j \mid x) \quad \forall j \neq i
\]  \hspace{1cm} (2-4)

which minimizes the Bayes risk for a symmetrical loss function (i.e. where the loss incurred from erroneous classification of \(x\) is equal to 1, and it is 0 when \(x\) is classified correctly) [Duda et al., 2001]. This approach is designed for the cases when \(x\) can be unambiguously determined. This is not so, however, in the cases of low quality patterns. A natural way of handling such cases is to reject them. However, there is another possibility that we strongly advocate in this dissertation. Instead of discarding low quality patterns, which is undesired when good quality samples are not available, we propose to introduce a notion of feature uncertainty. One can imagine an operational environment in which a decision is required and the low quality pattern is the only sample available. The goal of this thesis is to show that as long as the uncertainty can be quantified, the classification can still be performed in a meaningful way even in low quality...
cases. What uncertainty means is that there may be more than one plausible value of features that need to be examined. Chapter 4 contains a presentation of the orientation field estimation method where the low quality is expressed in a form of a variance. The estimates that are characterized by high variance are reinforced using the information from the neighborhood.

### 2.2 Quality Aware Classification

Numerous approaches presented in the literature on fingerprint recognition implicitly use the notion of quality but publications aimed explicitly at quality with applications to fingerprints are fairly recent [Chen et al., 2004], [Chen et al., 2005]. At a local level, within a window size comparable to a fingerprint ridge-valley cycle, a measure of local coherence of squared gradients [Kass and Witkin, 1987], [Bazen and Gerez, 2002], [Jiang, 2005] is used to quantify the reliability of an orientation estimate. The presence and strength of the pattern can be measured using gray scale intensity variance [Maltoni et al., 2003]. Fingerprint segmentation [Bazen and Gerez, 2001] can be accomplished using a combination of coherence of squared gradients, and intensity mean and variance. Similarly, quality measures can be used as indicators of which regions of an image require enhancement, or which regions are too degraded to be useful for processing and need to be rejected [Hong et al., 1998]. In [NIST, 2004] and [Chen et al., 2005] a scalar (global) quality metric is defined and is used to predict classification outcome. However, we argue that the treatment of the effect that quality has on subsequent processing stages has not been fully explored. In particular, it has not been shown how the quality of a raw pattern is propagated through the system and how it affects the final outcome - classification. This dissertation is an attempt to investigate how side information that reflects quality can be used to assist pattern processing and classification. Low quality is understood as a factor that introduces an uncertainty into the processing. The new approach presented in this work is based on the premise that features extracted from a given pattern are not known for certain but, on the contrary, they are assumed to represent a range of values.

The uncertainty is governed by stochastic processes. In the case of fingerprints there are numerous factors that have stochastic nature. A pattern can be collected in the presence of noise. It can be partial as a result of occlusion or fingerprint misalignment on a sensor. These processes are often too complex to be fully accounted for, but their effects manifest themselves in processing and classification. For instance, local ridge orientation estimation can be performed based on a maximum of an orientation signature. Such a signature is likely to have a number of maxima. Neglecting them, which is the case in the commonly accepted local ridge orientation estimation method based on the principal component analysis (PCA) of gradient field [Bazen and
Gerez, 2002], often leads to discarding useful information. In this work it is proposed that such information should be used and that using it can lead to performance improvement.

In the traditional approach one has a point estimate of a feature vector \( \mathbf{x} \) for which a posterior probability that it belongs to a class \( \omega_i \) can be calculated. The formulation of the classification problem using Bayesian probability does not explicitly take quality into account. It is acknowledged that there is an inherent variability in the problem. For \( N \) classes one seeks the class conditional probabilities \( p(\mathbf{x} | \omega_i) \) \( i = 1,2,...N \) and \( N \) class priors \( p(\omega_i) \) \( i = 1,2,...N \).

This is sufficient to construct a classifier that minimizes error due to (2-4). The variability inherent to the problem is captured by class conditional probabilities that are usually estimated from samples. The consequence of the inherent variability is a non-zero probability of an error. The error can be of two types: false acceptance – when a pattern is erroneously assigned to class \( \omega_i \), and false rejection – when a pattern from \( \omega_i \) is assigned to \( \omega_j \), for \( i \neq j \). Using the notation from the preceding section, the false rejection error probability for class \( \omega_i \) given the feature vector \( \mathbf{x} \) generated by \( \omega_i \) can be written as

\[
P_{\omega_i | \omega_i} = P(\mathbf{x} \notin R_i | \omega_i) = 1 - \int_{R_i} p(\mathbf{x} | \omega_i) p(\omega_i) d\mathbf{x}
\]

The above represents the minimum error if the feature space is partitioned into the regions \( R_i \) according to (2-4), i.e.

\[
\mathbf{x} \in R_i \quad \text{if} \quad p(\omega_i | \mathbf{x}) > p(\omega_j | \mathbf{x}) \quad \forall j \neq i
\]

For a given value of a feature vector \( \mathbf{x} \) from \( \omega_i \), the probability of misclassification equals

\[
P_{\omega_i | \omega_i} = 1 - p(\omega_i | \mathbf{x}) p(\mathbf{x})
\]

Here is how we propose that the knowledge of quality can be explicitly included in the classification. We show how the classifier can be modified to include information about quality and how the error rates are influenced. Let us refer to the generic diagram of a pattern recognition system from Figure 2.1 and indicate the processing steps where the quality can be assessed and the steps where it can be used.
The quality influences the ability of the feature extraction algorithm to estimate \( x \) from the input \( I \). It is difficult to discuss quality without a specific context. Therefore, here we indicate the conditional dependence of the feature vector \( x \) on the raw pattern \( I \). The notion of quality throughout this work is discussed in the context of particular preprocessing and feature extraction methods. Low quality of \( I \) is associated with greater uncertainty in the estimated quantity \( x \). At the preprocessing stage the quality information can be used to decide whether the pattern is suitable for classification or whether it needs to be discarded. The latter approach is traditionally used. The probability density \( p(x|I) \) denotes our uncertainty associated with the feature vector \( x \), where \( p(x|I)>0 \), for some \( x \in X' \subseteq X \). In the quality aware processing paradigm we require that the uncertainty associated with the estimate of the feature vector \( x \) is used by the classifier. The class conditional probabilities and class priors remain unchanged since we make an assumption that the structure of the classification problem is independent of the quality of collected samples. However, the usual rule that minimizes Bayes error, shown in (2-4), needs to be altered. One could present it in the graph shown in Figure 2.4, in which arrows show the causal ordering.

\[
\omega_i \rightarrow x \rightarrow I
\]

**Figure 2.4.** The sample \( I \) contains the quality information that can be included in classification. The class \( \omega_i \) generates features \( x \) that are embedded in the raw pattern \( I \), possibly distorted as a result of low quality.

In order to justify the statements above, let us use the example from the previous section. Consider the two cases: 1) the feature vector is assumed to be known with probability 1 as in traditional pattern classification approach, and 2) instead of a feature vector we have a probability...
density function \( p(x | I) \) that represents the uncertainty about the knowledge of the true value of \( x \). Such a situation can take place, for instance, when the quality of the raw pattern \( I \) is low. As can be seen from the Figure 2.1 and Figure 2.3, there is a number of processing steps prior to actual feature extraction. The uncertainty associated with the feature should therefore take into consideration the quality of the raw pattern \( I \), the way in which the preprocessing is influenced by the quality, and the properties of the processing operations. This requires a mechanism of conveying the information about quality to the classification stage, as indicated in Figure 2.3. To reflect this quality aware processing we can slightly modify the equation (2.3), by replacing the feature vector \( x \) by \( p(x | I) \):

\[
p(x | I) = F(p(I))
\]  \hspace{1cm} (2.8)

Based on the graph in Figure 2.4, we can write the posterior probability \( p(o | I) \)

\[
p(o | I) = \frac{p(I | o)P(o)}{p(I)} = \int \frac{p(I | o, x)p(x | o)P(o)dx}{p(I)}
\]

We can use the fact that \( I \) is conditionally independent of \( o \) given \( x \), i.e. \( p(I | o, x) = p(I | x) \).

\[
p(o | I) = \int \frac{p(I | x)p(x | o)P(o)dx}{p(I)} = \int \frac{p(x | I)p(I)p(x | o)P(o)dx}{p(x)} = \int p(x | I)P(o)dx = \int p(x | I)p(o | x)dx
\]

The classifier, after including the feature uncertainty \( p(x | I) \), can therefore be written as

\[
\text{choose } o \text{ if } \int p(o | x)p(x | I)dx > \int p(o_j | x)p(x | I)dx \quad \forall j \neq i
\]  \hspace{1cm} (2.9)

The probability of error (2.5) needs to be modified by including the feature uncertainty.
\[ P_{e,\omega_i} = P(I \notin S_i, \omega_i) = 1 - \int_{S_i} p(I, \omega_i) dI = 1 - \int_{S_i} \int p(x, I, \omega_i) dx dI = \tag{2-10} \]
\[ = 1 - \int_{S_i} p(\omega_i | x) p(x | I) p(I) dx dI \]

Note that both the classifier and the probability of error are influenced by the uncertainty of the feature via \( p(x|I) \). \( I \) is the pattern obtained at the input for which the probability \( p(I) \) is in realistic scenarios difficult to obtain. The consequence of including \( p(x|I) \) is that the posterior probability that \( x \) is generated by \( \omega_i \), \( p(\omega_i | I) \), is increased if the large portion of the probability mass of \( p(x|I) \) is in the region of the \( R_i \) associated with the class \( \omega_i \).

For the purpose of illustration, let us assume that the following \( p(x|I) \) has been obtained (Figure 2.5). \( p(x|I) \) can be, for example, a response of the two-dimensional anisotropic filter to an oriented pattern.

\[
p(x | I) = 0.3 \cdot N(\mu_{s1}, \Sigma_{s1}) + 0.7 \cdot N(\mu_{s2}, \Sigma_{s2})
\]

with the following parameters

\[
\begin{align*}
\mu_{s1} &= [-30 -20]^T \\
\Sigma_{s1} &= \begin{bmatrix} 25 & 0 \\ 0 & 25 \end{bmatrix} \\
\mu_{s2} &= [20 20]^T \\
\Sigma_{s2} &= \begin{bmatrix} 121 & 0 \\ 0 & 121 \end{bmatrix}
\end{align*}
\]

![Figure 2.5. The marginal probability \( p(x) \) - a), and \( p(x|I) \) - b).](image)
The posterior probabilities computed for the feature vector corresponding to the maximum of $p(x | I)$,

$$\hat{x} = \arg \max_x \{ p(x | I) \}$$ are equal to $p(\omega_1 | \hat{x}) = 0.0284$, $p(\omega_2 | \hat{x}) = 0.9716$.

Based on these results, we conclude that $\hat{x}$ was most likely generated by $\omega_2$. It is interesting to examine the outcome of classification performed according to eq. (2-9) in which the feature signature is used:

$$\int p(\omega_1 | x) p(x | I) dx = 0.6982$$

$$\int p(\omega_2 | x) p(x | I) dx = 0.3018.$$
CHAPTER 3

QUALITY OF DIGITAL FINGERPRINT IMAGES

The material presented in this chapter provides foundations for understanding the concept of fingerprint identification and categorization in low image quality scenarios. Section 3.1 introduces the notion of features and categories. In Section 3.2 important quality defects are identified followed by discussion of their properties in Section 3.3.

3.1 Salient Features

The pattern recognition principles discussed in the previous chapter apply to a variety of problems. One proceeds by first finding the set of features that allow distinguishing any two patterns of the given type and, second, by designing algorithms that allow for reliable feature extraction. Due to the abundance of applications, the focus of this work is on automated recognition of fingerprints, which are probably the most widely used biometric identifiers. Fingerprint images have been used in modern forensics for over 100 years; they are well examined and considered as a reliable means of identification. Automated systems for fingerprint recognition are available since the 1960's, but became popular only recently and now have made their way into consumer products mainly as means for access control in the automotive industry, handheld, portable and stationary electronics.

Extensive studies on fingerprint individuality are available, see [Maltoni et al., 2003] and references therein. Here we provide only an introduction that can facilitate the reading of the material presented in the sequel. Fingerprints share some features, most notably that the skin of the fingertip is characterized by a quasi-periodic pattern of ridges, which due to their function can be referred to as friction ridges. When a ridge-furrow impression is left on a smooth surface, it is referred to as a fingerprint. The periodicity of the ridge-furrow pattern is fairly constant among the population. It has been determined that the average periodicity is equal to about 500 μm [Maltoni, et al., 2003]. Within the class of fingerprint patterns one can distinguish a number of categories identified by global characteristics of ridge patterns. The task of classification to one of the categories can be referred to as fingerprint categorization in order to distinguish it from matching or identification. Fingerprint categorization is instrumental in facilitating search of large
fingerprint databases. The scheme accepted by most law enforcement agencies uses Galton-Henry classes [Maltoni et al., 2003]; the five most common are arch, tented arch, left loop, right loop and whorl – shown in Figure 3.1.

Figure 3.1. The most common fingerprint classes NIST DB4\(^1\); arch a), tented arch b), left loop c), right loop d), whorl e).

The proportions of the fingerprint categories in the general population are known: arch 3.7%, tented arch – 2.9%, left loop – 33.8%, right loop – 31.7% and whorl – 27.9%, [Maltoni et al., 2003]. It justifies the indexing scheme based on these categories. For instance, identifying the category of the fingerprint as an arch limits the search size of the database to 37 out of each 1000 fingerprints, which is clearly advantageous. One needs to keep in mind the probability of categorization errors – both those inherent to the problem and those due to low image quality.

Within ridge-valley patterns one can identify so-called SPs. These are points of high curvature and can be of two types: core or delta. More details can be found in [Maltoni et al., 2003] and the references therein. The illustrations in Figure 3.2 show examples with marked singularities. SP counts are sufficient for the 3-class categorization task. It has been determined that an arch contains no SPs, a tented arch, left and right loops contain one core and one delta, and a whorl contains two core SPs and 2 delta SPs. If one includes information about the relative position of core and delta SPs and the number of ridges that a line connecting the core and delta SPs crosses,

\(^1\) NIST DB4 – NIST database is an accepted standard for testing fingerprint recognition systems.
then it is possible to distinguish left loop from right loop and a tented arch from loops. SPs alone are typically not reliable enough. Often as a result of low quality, they can be left undetected or, as it is most often the case with delta SP’s, they are left outside of the impression. For this reason, and also in order to allow for categorization into more than the basic three classes, more features are often used. They can include ridge line properties, counts of ridges between SP’s or features of the fingerprint ridge-furrow orientation field.

![Image of fingerprint impressions with singular points](image)

Figure 3.2. Core and delta singular points NIST DB4; tented arch with core and delta SPs a), left loop b), right loop c) and whorl d) with two core SPs (○) at the center and two delta SPs (▲) on the periphery of the impression.

In fact, a number of alternative categorization schemes have been developed with the goal of providing automated indexing of large databases with low error rates. These applications involve features such as ridge lines, pseudo-ridge lines, orientation field and SPs. The categories do not need to be equivalent to Galton-Henry classes as long as they can be identified reliably. A summary of the approaches following this notion with examples of algorithms can be found in [Maltoni et al., 2003]. For the reference on error rates in biometric systems in which indexing scheme is applied refer to [Wayman, 1999, 2005].

From the standpoint of identification (one-to-many matching) or authentication (one-to-one matching), more detailed features are required. It has been established that fingerprint patterns provide enough degrees of freedom to ensure that the occurrence of two identical fingerprints is improbable. Depending on the application, the requirement on the type and the number of features required in order to establish a match, or lack of thereof, can vary. In identification tasks
the requirements are more stringent — more features and more robust techniques are required due to the number of comparisons. This is in contrast to authentication where the identity of the fingerprint holder is assumed to be known \textit{a priori} and needs only to be verified. In both cases the features used in matching include fine details of the ridge-valley pattern — so-called minutiae points. The table below shows the most common minutiae types and illustrating examples [Maltoni et al., 2003].

\begin{table}[h]
\centering
\caption{The most common minutiae [Maltoni et al., 2003]}
\begin{tabular}{|c|c|c|c|}
\hline
Minutiae & Type & ANSI taxonomy & FBI minutiae coordinate model \\
\hline
 & termination & \checkmark & \checkmark \\
\hline
 & bifurcation & \checkmark & \checkmark \\
\hline
 & lake &  &  \\
\hline
 & independent ridge &  &  \\
\hline
 & point or island &  &  \\
\hline
 & spur &  & \checkmark \\
\hline
 & crossover &  &  \\
\hline
\end{tabular}
\end{table}

According to the ANSI taxonomy, all minutiae that do not fall under the three categories highlighted in the third column are classified as undetermined. In practice the most common minutiae, bifurcations and terminations, are also the most reliable to detect automatically and have been employed in the FBI coordinate system.

In order to perform matching of two fingerprints, one is required to know the location, orientation and type of minutiae. Legal standards define a minimum number of matching minutiae points that are required, which vary from country to country. However, professional fingerprint examiners use more than just minutiae points. Details of ridge lines, their shape, and the number of ridges between minutiae are often used to reinforce matching. [Maltoni et al., 2003] give the following
factors that need to be evaluated in the process of examination: 1) the same global pattern – arch, loop, whorl, etc., 2) correspondence of the minutiae pairs in compared fingerprints, 3) minimum number of corresponding minutiae found, 4) "corresponding minutiae details which must be identically interrelated". In order to be equally efficient as human examiners, automatic fingerprint matching algorithms need to follow similar procedures. This task, however, is difficult to accomplish for obvious practical reasons.

It is believed that even more discriminatory information is embedded in so-called Level 3 features. In a recent publication [Jain et al., 2007] give the following definitions: 1) Level 1 features define pattern type or category of the fingerprint. They are defined by the global characteristics of ridges as arch, loop, whorl etc., 2) Level 2 features comprise minutiae points; most common among them are ridge termination and ridge bifurcation, and Level 3 features include even finer details of the ridge lines such as their contours, shape and location of sweat pores, and other permanent features that can have the appearance of creases and scars.

The categorization task (according to the global ridge pattern type) faces the problem of large intra-class variability. Loop patterns, for example, can be characterized by different numbers of ridge lines between delta and core SPs, different ridge curvature, and differences in fine details including minutiae or sweat pores. On the other hand, elastic deformations of fingerprints do not usually affect the global ridge pattern significantly. This, however, is an important factor in the case of matching of individual fingerprints. Elastic deformations of the fingerprint are created as a flat impression of a three dimensional object is pressed against a surface. They change relative distances between minutiae and affect their relative orientations. If the non-elastic transformation could be somehow available and reversible, then the fingerprint matching could be simply accomplished by measuring pixel-wise distance between two images. This is not possible in practical applications. Fortunately, there are techniques that facilitate the matching process in the presence of deformations; for a more detailed treatment we refer the reader to [Maltoni et al., 2003] and numerous references provided therein.

3.2 Defects – Qualitative Approach

In Chapter 2 it has been shown that uncertainty of the features can be used in the classification process. It is interesting to investigate how the uncertainty arises in typical operational environment. Here we focus on the local ridge orientation and orientation field estimation and the uncertainty due to the low quality of the ridge-furrow pattern. Examples of fingerprint images are
presented in Figures 3.3 through 3.10. They contain common quality defects and can serve as a starting point for the development of defect models. The goal is to develop a model for the quality defects that can be used in analysis of their impact on feature estimation. Of particular interest are defects that can pose problems in classification tasks by introducing spurious features or by precluding detection of legitimate features. Local ridge orientation features are typically used in the fingerprint categorization task that is used as a case study in this work. The first difficulty that one faces is how to tell what may be an imaging defect from what is the feature of the fingerprint itself. The examples are scars and creases—some of them can be permanent and therefore can be used in the classification, while others can originate as imaging artifacts and should therefore be ignored. It is known that the fingerprint ridge pattern is continuous and that local ridge orientation changes in a smooth fashion with the exception of SPs. Those characteristics can be used as a guide in finding suitable defect models. Scars and creases introduce discontinuity to the ridge pattern and distort the local orientation field, which would otherwise be smooth. From this point of view, creases and scars can be treated as quality defects. Sensing technology can introduce specific defects such as periodic patterns, noise, mosaic artifacts, over-emphasized ridges for moist skin condition and under-emphasized ridges for dry fingerprints. In particular, different sensors have different abilities to reproduce the original fingerprint. For example, potentially very high fidelity three-dimensional images, including the surface of ridges as well as valleys, can be obtained using ultrasonic sensors. On the other hand, capacitive sensors are typically unable to provide details of the surface of the furrows. The latter is, however, of minor concern as the ridges contain sufficient discriminatory information.

Some examples will enable the reader to better appreciate the nature of different defects. The images are part of the test sets collected for the purposes of the Fingerprint Verification Competition – FVC 2000, FVC 2002 [Maltoni et al., 2003], FVC 2004 [Cappelli et al., 2006] test sets. The defect numbers have been added to the figures for easier reading, and they reflect the numbering in the Table 3.2 that summarizes the observations. They are presented in the following order: 1) global and slow intensity variations, 2) local intensity variations along the ridge lines, 3) overemphasized ridges, 4) anisotropic defects – bright, 5) anisotropic defects – dark, 6) low contrast, 7) low signal-to-noise ratio, and 8) other defects. The ordering does not reflect the relative frequency nor the importance of the defects.

The images in Figure 3.3 illustrate the effect of gradual intensity variation across the image. It is frequently present in fingerprint images and it is typically associated with contrast variation. In these images, it can be observed that the ridges are darker towards the image centers—it can be
verified that the mean intensity increases towards the boundary of the fingerprint silhouette. The intensity variations in the examples presented in Figure 3.3 have insignificant effect on ridge orientation estimation. This is only true for intensity variations that have global character and do not exhibit high-rate local variations.

![Figure 3.3. Images from FVC2000 DB1b (a,b) and FVC2002 DB2b (c,d) – defect 1.](image)

In the Figure 3.4 below, the example images are characterized by significant local intensity variations along ridge lines. Discontinuities and gaps in ridges are present. This situation can occur if insufficient pressure is applied onto the sensing surface, or it can be due to the skin conditions, where dry skin has been reported to produce similar effects. Orientation estimation can be affected if the variations along ridge lines are significant. Typically, as in the examples shown below that are not severely corrupted, one can take advantage of the bandpass character of the ridge-valley pattern and separate the noisy artifacts from the pattern by using a matched filter. The filter is tuned to local ridge orientation and frequency. This method is more robust when the analysis window spans multiple ridges but, unfortunately, the size of the window can not be made arbitrarily large as the local orientation can be only assumed constant locally.

![Figure 3.4. Synthetic images from FVC2000 DB4b (a,b) and live scans FVC2004 DB1b (c,d) - defect 2.](image)
The examples of the images with overemphasized ridges are shown below in Figure 3.5. It can be seen that in this case the ridge lines are wider and neighboring ridges tend to merge. This can lead to decreased contrast and in extreme situations ridges cannot be discerned. As it will be shown further, overemphasized ridges are characterized by the presence of higher order harmonics of the ridge-furrow frequency and the energy of the signal is dispersed across the spatial-frequency spectrum.

![Figure 3.5. Images from FVC2000 DB2b (a,b) and FVC2004 DB1b (c,d) - defect 3.](image)

Figure 3.5 contains examples of images in which the ridge-furrow pattern is crossed by oriented patterns, likely originated as skin folds. Their effect on the orientation estimate depends on their spatial extent and manifests itself as bias in the orientation estimation in the case of mean squared error (MSE) estimators (see Chapter 4). In the case of methods in which the maximum of the orientation signature is used, the orientation of the oriented defect can be mistaken for the actual ridge orientation.

![Figure 3.6. Images from FVC2000 DB1b (a,b) and FVC2000 DB3b (c,d) - defect 4.](image)

Images presented in the Figure 3.7 contain a similar type of defects as images in Figure 3.6. However, they are characterized by dark intensity levels, close in value to intensities of ridges. In
the images shown below they appear to be associated with skin folds, imaging artifacts (the two images on the left) and scar tissue (the image on the right).

![Images](image1.png)

Figure 3.7. Images from FVC2004 DB3b (a,b), and NIST special database 4 (c) - defect 5.

The low contrast images shown below in Figure 3.8 present a challenge to the segmentation process as the fingerprint ridge pattern is often difficult to distinguish from the background. The low contrast images often can be characterized by significant variations of the intensities along the ridges and often have low signal-to-noise ratios, which in turn implies increased variance of orientation estimates. These images show that the contrast range can be limited to gray end (intensities closer to zero) or white end of intensity range, but often there are regions of low contrast at different intensity intervals as it can be seen in the case of the image b) in Figure 3.8.

![Images](image2.png)

Figure 3.8. Images from FVC2000 DB1b (a,b), DB2b (c), DB2b (d) - defect 6.

Figure 3.9 contains fingerprints acquired with low signal-to-noise ratio. It should be emphasized that although the cases of low SNR are often associated with low contrast, the two cases should be distinguished from one another. The particularly low quality can be observed in the image c).
The examples shown above by no means characterize all possible types of quality defects that can occur in fingerprint images. The author’s intention is to provide as complete an overview as possible based on the available databases. The rare, but nevertheless possible, cases of other defect types may involve missing regions of ridge pattern (Figure 3.10 a,b,d), regular patterns resulting from sensing technology (Figure 3.10 c) and possibly more.

![Figure 3.9. Images from FVC2000 DB2b (a), DB3b (b), FVC2002 Db3b (c) - defect 7.](image)

The following table (Table 3.2) contains the summary of defects identified based on the examples presented above. Probable root causes of the defects are listed in the rightmost column. The analysis of the impact the defects have on selected processing techniques is conducted in Chapter 6. The table below has been prepared based on an overview of the literature and after examining a large number of images collected using live sensors [Cappelli et al., 2006] and rolled impressions from NIST special database 4.

![Figure 3.10. Images from FVC2000 DB1b (a), FVC2002 DB3b (b), FVC2004 DB1b (c) - defect 8.](image)
### Table 3.2. Fingerprint defects.

<table>
<thead>
<tr>
<th>No.</th>
<th>Defect type</th>
<th>Probable root cause</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Slow average intensity variations across fingerprint image, with low local intensity variations along ridge lines</td>
<td>Uneven pressure, skin conditions</td>
</tr>
<tr>
<td>2</td>
<td>Significant intensity variations along ridge lines/ under-emphasized ridges</td>
<td>Insufficient pressure, dry skin conditions</td>
</tr>
<tr>
<td>3</td>
<td>Over-emphasized ridges</td>
<td>Excessive force exerted on a sensor, moist skin conditions</td>
</tr>
<tr>
<td>4</td>
<td>Anisotropic artifacts, characterized by intensity levels close to valley regions (usually bright)</td>
<td>Skin creases, scars</td>
</tr>
<tr>
<td>5</td>
<td>Anisotropic artifacts, characterized by intensity levels close to ridge regions (usually dark)</td>
<td>Scars, skin creases, sensing artifacts</td>
</tr>
<tr>
<td>6</td>
<td>Low contrast</td>
<td>Insufficient pressure/ dry skin conditions or too much pressure/ moist skin conditions</td>
</tr>
<tr>
<td>7</td>
<td>Low signal to noise levels</td>
<td>Insufficient pressure or dry skin conditions</td>
</tr>
<tr>
<td>8</td>
<td>Other – missing region of ridge pattern, distorted ridge pattern, presence of a regular pattern across the image</td>
<td>Scars, excessive pressure on the sensor, lack of contact between skin and the sensor, sensor artifact or mosaic artifact</td>
</tr>
</tbody>
</table>

### 3.3 Models of Quality Defects

A valuable insight into the effect of defects in the fingerprint images on feature extraction can be gained by studying the defects themselves. One way of achieving this is through modeling. Ideally, the models should reflect the processes that give rise to quality defects. As this can lead to rather complex solutions, we content ourselves with models that give rise to similar phenomena as their real world antecedents, and that are simple enough to permit a certain degree of analysis.

The publications on the topic of defect models for fingerprint images are limited. The only known results of defect modeling have been published in the context of synthetic fingerprint generation by Cappelli, Maio and Maltoni [Cappelli et al., 2000, 2002, 2004a, 2004b]. The ability to generate large numbers of realistic images is attractive from the perspective of verification of recognition systems. At present the verification is most reliably conducted using databases of natural images. The process of collecting of the images is time consuming and is accompanied by issues of privacy. Yet another factor is associated with the choice of test samples, especially important in small-sized datasets. The test databases are required to be challenging for the state of the art systems (not too difficult and not too easy). There are a number of databases available for...
performance testing, but they provide a limited variability with respect to presence of quality defects. Due to the complexity of the task, the exhaustive approach is not practical or even not possible. The synthetic generators have certain shortcomings that preclude them from being widely accepted for the purpose of performance rating. [Maltoni et al., 2003] provides results of FVC - an initiative geared towards fingerprint systems testing where synthetic databases are included in the test set. It can be expected that eventually the synthetic pattern generators will be able to produce patterns that will mimic the real ones closely enough. The synthetic set used in the FVC competition has been generated using the SFINGE generator developed at the University of Bologna, Italy. The approach presented here differs from the one used in SFINGE and is conceived to facilitate a systematic approach to evaluation of the effects that particular defects have on pattern classification. In the case of SFINGE, on the other hand, the goal is to ensure natural appearance of the synthetic fingerprints. Briefly, the approach in SFINGE is to first generate a binary pattern of ridges and valleys (a master fingerprint pattern) and then alter it to ensure a more realistic appearance. The altering of the binary pattern of ridges and valleys can be viewed as defect generation. It is accomplished by adding noise and applying morphological operators (erosion and dilation) to alter the ridge widths. The ridge width variations simulate the natural phenomena associated with varying pressure and skin conditions. An averaging operator is applied to smooth the image. Other operators that emulate variable positioning of the fingerprint on the surface and elastic distortion are also implemented.

The purpose of this section is to propose a defect model that is amenable to analysis and yet provides sufficient variability reflecting that of natural patterns. The observations, some of which have been presented in the preceding part of this thesis, indicate that sufficient variability can be modeled by additive and multiplicative components. The additive part comprises noise and fine scale artifacts. Noise, such as a Gaussian noise, can originate at the sensing stage. Fine scale artifacts in this work are modeled as Gaussian functions. They can be used to simulate the irregularities of the image of the skin surface and can be viewed as additive or multiplicative. In the case of SFINGE, on the other hand, the noise is simulated via binary white blobs of different sizes and shapes that are added to the image. Since the intensity range in digital images is finite, the intensities that exceed the maximum threshold are truncated and this effectively introduces a multiplicative component. The resulting noisy image is then smoothed. In this work the Gaussian functions are combined to form a multiplicative mask according to the formal description presented below.
For simplicity, in this work the ideal pattern emulating the fingerprint is assumed to have the form of a sine wave. Traditionally in fingerprint images ridges are represented as dark and valleys as bright, and the images shown in this thesis adhere to this representation. To simplify the design of the multiplicative defect mask, however, an inverted signal (pattern of ridges) and an inverted mask are conveniently introduced and are identified by the subscript "inv" for clarity. Let us assume that the image is represented in a normalized form in which intensities are taken from the interval \([0,1]\), with 1 corresponding to white and 0 to black. Consequently, the inverted image will have bright ridges and dark valleys. In the case of the inverted mask, zero denotes lack of signal, i.e., the defect obscures the pattern completely, and the all-ones mask signifies no distortions.

Let \(s(x,y)\) be a two-dimensional signal - a real non-negative valued function defined on a Cartesian grid. Let us further assume that the defect model can be defined as a multiplicative mask \(m(x,y)\) and that at each point \((x,y)\) an additive noise \(n(x,y)\) with known statistical properties is present. The signal \(s(x,y)\) modified this way can then be written as

\[
s_m(x,y) = 1 - s_{inv}(x,y) \cdot m_{inv}(x,y) + n(x,y)
\]  

(3-1)

where \(s_{inv}(x,y)\) and \(m_{inv}(x,y)\) denote the inverted signal and inverted mask.

\[
s_{inv}(x,y) = (1 - s(x,y))
\]

\[
m_{inv}(x,y) = (1 - m(x,y))
\]

(3-2)

After substituting (3-2) into (3-1) and completing the multiplication, we have

\[
s_m(x,y) = 1 - (1 - s(x,y)) \cdot (1 - m(x,y)) + n(x,y)
\]

\[
= s(x,y) + m(x,y) - s(x,y) \cdot m(x,y) + n(x,y)
\]

\[
= s(x,y) + m(x,y) \cdot (1 - s(x,y)) + n(x,y)
\]

(3-3)

The result of application of the defect model described above by (3-3) has been shown in Figure 3.11. The synthetic images consist of a 2-D sine wave altered by the multiplicative mask \(m(x,y)\) consisting of a superposition of Gaussians with varying scale parameters.
Figure 3.11. Sample images $s_m(x,y)$; underlying pattern is a sine wave and the defects are modeled by a multiplicative mask composed of a superposition of Gaussian functions.

In fingerprint images one frequently encounters oriented defects that cross a number of ridge lines and that are typically characterized by a regular, very often linear, shape. Such defects can conveniently be modeled using a multiplicative mask $m(x,y)$. Rectangular and ellipsoidal shapes have been tried in this work as oriented defect models, but the resulting images had artificial appearance. The similar effect was produced by masks created using Gaussian functions with non-diagonal covariance matrices. It has been determined that a good choice for a defect mask is a superposition of circularly symmetric Gaussians. The superposition of such masks or kernels allows for flexibility in generating defect patterns resembling isotropic and anisotropic shapes found in natural fingerprint patterns.

The superposition of elementary defect models can be written as

$$m(x, y) = \sum_{(\mu, \sigma)} f_{\mu, \sigma}(x, y)$$ \hspace{1cm} (3-4)$$

where the 2-element vector $\mu$ represents the translation.
\[
\mu = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}
\]

and the scalar \( \sigma \) represents the scale. For the case when \( f \) is a circularly symmetric Gaussian

\[
f_{\mu,\sigma}(x, y) = \exp\left( -\frac{(x - \mu_x)^2 + (y - \mu_y)^2}{2\sigma^2} \right)
\]

The parameters \( \mu_x, \mu_y, \sigma \) can be modeled as deterministic or random, depending on the modeled defect type. Figure 3.12 b) depicts the result of applying a mask \( m(x,y) \) to a sinusoidal pattern. The mask \( m(x,y) \) includes an anisotropic defect model shown in Figure 3.12 a).

Further modification of the model is needed in order to account for the loss of information in situations when the signal is either not present or causes the saturation of the sensing device. In such cases it can be assumed that only an additive noise is present.

Let the relationship between the input and output be defined as the following relationship

\[
m_s(x, y) = \begin{cases} 
\frac{m(x, y)}{c} & m(x, y) \leq c \\
1 & m(x, y) > c
\end{cases}
\]

\[m(x, y), c \in (0,1)\]
where $c$ is a threshold defining the cutoff level. When $c=1$, no saturation takes place, and (3-7) should not be used. For the case when $c=1$, assume $m_s(x, y) = m(x, y)$. The saturation function can be implemented using the Heaviside step function $h(t)$. The new mask that takes saturation into account can then be written as

$$m_s(x, y) = \frac{1}{c} m(x, y) - \left( \frac{1}{c} m(x, y) - 1 \right) \cdot h\left( \frac{1}{c} m(x, y) - 1 \right)$$

(3-7)

where $m_s(x, y)$ denotes the mask after applying the saturation function,

$$h(t) = \begin{cases} 
1 & t > 0 \\
0 & \text{otherwise} 
\end{cases}$$

(3-8)

and consequently

$$s_m(x, y) = s(x, y) + m_s(x, y) \cdot (1 - s(x, y)) + n(x, y)$$

(3-9)

The effect of reduced contrast can be achieved using a linear mapping as follows:

$$s_{cr}(x, y) = c_{cr}(x, y) \cdot s_m(x, y) + d_{cr}(x, y)$$

(3-10)

$c_{cr}(x, y) \in [0, 1]$

d_{cr}(x, y) \in [0, 1 - c_{cr}(x, y)]$

The contrast reduction coefficient $c_{cr}$ defines the degree of compression and $d_{cr}$ can be used to control the shift of the mean intensity. The range of $d_{cr}$ is specified in order to ensure that the resulting intensities are not greater than 1. The parameters $d_{cr}$ and $c_{cr}$ can otherwise be specified independently. It can be verified that for $c_{cr}=1$, the contrast remains unchanged; the case of $c_{cr}=0$ denotes absence of the pattern. The examples showing the effect of contrast reduction are presented in Figure 3.13 below.
The defect resulting from excessive or insufficient pressure applied to the sensor surface can be modeled as a nonlinear mapping governed by the pressure coefficient $\alpha_{pr}$

$$s_{pr}(x, y) = (s_m(x, y))^{\alpha_{pr}(x, y)}$$

$$\alpha_{pr}(x, y) \geq 0$$

Setting $\alpha_{pr}(x, y) > 1$ has the effect of ridge thinning; increasing $\alpha_{pr}(x, y)$ reflects increasing the pressure. For $\alpha_{pr}(x, y) \in [0, 1]$, and as $\alpha_{pr}$ gets smaller, the ridges become more emphasized. This effect can be combined with the saturation function to represent the cases in which the signal is lost as a result of the limited dynamic range of the sensor.

The relationship between $\alpha_{pr}$ and the ridge-line width can be found. In order to do so, first let us define the ridge width as the distance between opposite edges of the ridge at the intensity level equal to half the difference between top of the ridge and the bottom of the valley. For the signal $s_{pr}(x)$ defined as
\[ s_{\nu}(x) = (0.5 + 0.5 \sin(2\pi f_0 x))^\nu \]  
(3-12)

The width measured at \( s_{\nu}(x) = 0.5 \) leads to the following equation (3-13):

\[ \sin(2\pi f_0 x) = 2^{1-1/\alpha_{\nu}} - 1 \]  
(3-13)

The graph of (3-13) is illustrated in Figure 3.15 below, where the horizontal lines correspond to the levels at which different ridge widths can be found for different \( \alpha_{\nu} \).

![Graph](image)

**Figure 3.15.** The plot of \( \sin(2\pi f_0 x) \); the widths of the ridges can be found by measuring the length of the line segment between the intersections of the horizontal line with the sine function, according to (3-13); \( n \) denotes a sample number.

The ridge width is then given by the following formula:

\[ \Delta w = \frac{1}{2f_0} \cdot \frac{\arcsin(2^{1-1/\alpha_{\nu}} - 1)}{\pi f_0} \]  
(3-14)

For \( \alpha_{\nu} < 1 \), the second term on the right hand side of (3-14) is negative, thus indicating that the width of the ridge is greater than \( 0.5/f_0 \). For \( \alpha_{\nu} = 1 \) and \( \alpha_{\nu} > 1 \), the width of a ridge is unchanged and narrower than \( 0.5/f_0 \), respectively.
To summarize, we have proposed a defect model, which has the capability of emulating defects, listed in Table 3.2, according to the following strategy. Given the signal \( s(x,y) \), a sine wave:

1) Generate defect mask \( m(x,y) \)
2) Apply saturation operator to obtain \( m_s(x,y) \)
3) Apply mask \( m_s(x,y) \) to \( s(x,y) \) to obtain
\[
\begin{align*}
    s_m(x, y) &= s(x, y) + m_s(x, y) \cdot (1 - s(x, y)) \\
    s_p(x, y) &= s_m(x, y) \cdot f^p
\end{align*}
\]
4) Simulate the effect of pressure
\[
    s'(x, y) = s_p(x, y)
\]
5) Change contrast and mean intensity
\[
    s_c(x, y) = c_{cr}(x, y) \cdot s_p(x, y) + d_{cr}(x, y)
\]
6) Add noise
\[
    s'(x, y) = s_c(x, y) + n(x, y)
\]

Some of the properties of the defect model are presented in the following section, and a more detailed treatment of the effect of the modeled defects on orientation estimation is presented in Chapter 6.

### 3.4 Defects - Quantitative Approach

In this section, some of the properties of the defect models are discussed. We begin with the multiplicative mask \( m(x,y) \), which is a superposition of circularly symmetric Gaussian functions specified by different translations and different scale parameters. In particular, one can distinguish 1) large scale components that encompass the entire fingerprint pattern and 2) localized defects modeled by fine-scale Gaussians. In the case of ridge orientation features, it is crucial to examine orientations along which the spatial signal shows the greatest variation. As demonstrated in Chapter 6, orientation estimation methods are designed to find either orientation of maximum energy or MSE estimate. The accepted approach to orientation estimation, frequently cited in the fingerprint processing literature, is based on PCA of intensity gradients. The effect of the defect mask \( m(x,y) \) on gradients is insignificant given the condition that \( m(x,y)<c \) is satisfied and \( \sigma \) is large (\( c \) is the saturation threshold and for \( m(x,y)>c \) the effect of discontinuity needs to be accounted for). The gradient of the circularly symmetric Gaussian has the following form:

\[
    g(x, y) = \left[ \frac{df_{\mu,\sigma}(x, y)}{dx} \quad \frac{df_{\mu,\sigma}(x, y)}{dy} \right]^T = \frac{1}{\sigma^2} \exp \left( -\frac{(x-\mu_x)^2+(y-\mu_y)^2}{2\sigma^2} \right) \begin{bmatrix} x-\mu_x \\ y-\mu_y \end{bmatrix} \quad (3.15)
\]
The gradient magnitude is bounded from above by (3-16), the maximum of (3-15) which is located at radial distance σ from the mean, and is expressed as

\[ \|g(x,y)\| \leq \|g_{\text{max}}\| = \frac{1}{\sigma \sqrt{e}} \]  

(3-16)

The figure below shows \(\|g_{\text{max}}\|\) as a function of the scale parameter σ, Figure 3.16 a), and the intensity plot of \(\|g\|\) for a given σ, Figure 3.16 b). White denotes maximum magnitude. The gradient response is high for low values of scale parameter (σ→0). In practice, in the case of digital images, the exact value of \(\|g\|\) and the upper bound \(\|g_{\text{max}}\|\) depend on the type of gradient operator used and on the representation of intensities.

![Figure 3.16](image)

Figure 3.16. Maximum of local gradient of the f(x,y) in a (log-log) plot of scale parameter σ a), and \(\|g_{\text{max}}\|\) b), where the maximum of the gradient is located in r=σ from the center.

It can be shown that the fine-scale Gaussians have significant effect on the frequency response as well (see Chapter 6).

The gradient of the multiplicative defect mask \(m(x,y)\) is shown below. We used the fact that gradient is a linear operator and \(m(x,y)\) consists of a superposition of Gaussians indexed here using i.
\[ g(x, y) = \left[ \frac{dm(x, y)}{dx}, \frac{dm(x, y)}{dy} \right]^T = \sum_i \frac{1}{\sigma_i^2} \exp \left( -\frac{||d_i||^2}{2\sigma_i^2} \right) d_i = \sum_i g_i(x, y) \] (3-17)

where

\[ d_i = \left[ d_{x,i}, d_{y,i} \right]^T = \left[ x - \mu_{x,i}, y - \mu_{y,i} \right]^T \]

\[ \mu_i = \left[ \mu_{x,i}, \mu_{y,i} \right]^T \]

The sum in (3-17) does not need to include all the Gaussians but only those that are sufficiently close to \((x,y)\) as the gradient magnitude contributed by \(i\)th component decays quickly with the distance \(d_i\) from the mean \(\mu_i\). It can be verified using the following equation:

\[ ||g_i(x, y)|| = \frac{1}{\sigma_i^2} \exp \left( -\frac{||d_i||^2}{2\sigma_i^2} \right) \|d_i\| \] (3-18)

Substituting \(\sigma_i\) for \(||d_i||\) above leads to (3-16). The effect of information loss modeled by the formula (3-7) is difficult to illustrate for the entire mask \(m(x, y)\). In the regions of the mask that exceed the threshold \(c\), the gradient is zero. In the case of a Gaussian, (3-5), and given the cutoff threshold \(c\), (3-6), the radius within which the gradient is 0 is equal to

\[ ||d_i|| = \frac{\sqrt{\ln \left( \frac{0.5}{c} \right)}}{\sigma_i} \] (3-19)

The following is the illustration of the defect manifestation in the spectral domain. In order to keep the example simple, let us assume again that the signal has the form of a sine wave

\[ s(x, y) = 0.5 + 0.5 \sin(2\pi f_o x') \] (3-20)

where

\[ x' = x \cos(\theta) + y \sin(\theta) \] (3-21)

\[ u_o = 2\pi f_o \cos(\theta), \quad v_o = 2\pi f_o \sin(\theta) \]
with the Fourier Transform

\[ S(u,v) = 0.5 \delta (u, v) - j 0.25 \{ \delta (u - u_o, v - v_o) - \delta (u + u_o, v + v_o) \}. \quad (3-22) \]

The signal \( s(x,y) \) modified using the mask \( m(x,y) \) can be obtained after substituting (3-20) into (3-3), which gives

\[ s_m(x, y) = 0.5 + 0.5 \sin (u_o x + v_o x) + m(x, y)(0.5 - 0.5 \sin (u_o x + v_o x)) + n(x, y) \quad (3-23) \]

Assuming the Fourier Transform of \( m(x,y) \) is known and is equal to \( M(u,v) \), the Fourier Transform of \( s_m(x,y) \) can be written as

\[ S_m(u,v) = 0.5 \delta (u, v) - j 0.25 \{ \delta (u - u_o, v - v_o) - \delta (u + u_o, v + v_o) \} + 0.5 M(u,v) + j 0.25 (M(u - u_o, v - v_o) - M(u + u_o, v + v_o)) + N(u,v) \quad (3-24) \]

The result has been presented in the following Figure 3.17.

![Figure 3.17](image)

**Figure 3.17.** Synthetic image showing a sine wave distorted according to equation (3-23) a), and an amplitude spectrum without DC component b).

It can be seen that the signal's frequency component is dominant. The effect of the additive noise (Gaussian zero-mean, standard deviation 0.01) is insignificant. The influence of the multiplicative defect mask is visible at spatial frequency \((0,0)\) and at \((u_o, v_o)\) and \((-u_o, -v_o)\); it is mainly due to the horizontal artifact visible in the lower part of the synthetic image.
Local orientation estimation can be biased by oriented artifacts. The case of a single oriented artifact modeled by a superposition of Gaussians is presented below. For simplicity we assume that circularly symmetric Gaussians are distributed along a line.

Suppose that the multiplicative mask \( m(x,y) \) has the following form

\[
m(x,y) = \sum_i \exp \left( - \frac{(x - \mu_{x,i})^2 + (y - \mu_{y,i})^2}{2\sigma_i^2} \right)
\] (3-25)

where, for simplicity \( \sigma_i = \sigma \). Its Fourier Transform can be written as

\[
M(u,v) = 2\pi\sigma^2 \exp \left( - 0.5 \cdot (u^2 + v^2) \cdot \sigma^2 \right) \cdot \sum_m \exp(j\mu_{x,m}u + j\mu_{y,m}v)
\] (3-26)

Without loss of generality, for any linear defect it can be assumed that \( \mu_{y,m} = 0 \), and thus

\[
M(u,v) = 2\pi\sigma^2 \exp \left( - 0.5 \cdot (u^2 + v^2) \cdot \sigma^2 \right) \cdot \sum_m \exp \left( j\mu_{x,m}u \right)
\] (3-27)

It can be seen that the part under the summation is the most important; it represents the phase that in turn determines the shape of the frequency response. Under the assumption that the Gaussians are distributed along a line at sufficiently small intervals, it can be verified experimentally that as the ratio

\[
\frac{\max(\mu_{x,m}) - \min(\mu_{x,m})}{\sigma}
\]

increases, the shape of \(|M(u,v)|\) becomes more elongated along the v direction and narrower along the u direction.

The following defect model involves the pressure exponent \( \alpha \) used to control the width of the ridges and valleys. Its effect on the signal and on its gradient can be illustrated using the Fourier domain. For the signal \( s(x,y) \) distorted using simulated pressure, according to (3-11), and, after substituting (3-20) into (3-11) and after algebraic manipulations, we have that for an integer \( \alpha \).
\[ s_{pr}(x, y) = \left(0.5 + 0.5 \sin(u_ox + v_oy)\right)^\alpha = \sum_{k=0}^{\alpha} \binom{\alpha}{k} 0.5^k \sin^k(u_ox + v_oy) \]

\[ = \sum_{k=0}^{\alpha} \binom{\alpha}{k} 0.5^k \sum_{m=0}^{k} (2j)^{-k} (-1)^m \binom{k}{m} \exp(j(u_ox + v_oy)(k - 2m)) \]

where

\[ \sin^k(u_ox + v_oy) = \sum_{m=0}^{k} (2j)^{-k} (-1)^m \binom{k}{m} \exp(j(u_ox + v_oy)(k - 2m)) \]

From (3-28) the gradient can easily be computed

\[ \nabla s_{pr}(x, y) = \left[ \begin{array}{c} ju_ox \sum_{k=0}^{\alpha} \binom{\alpha}{k} 0.5^k \sum_{m=0}^{k} (2j)^{-k} (-1)^m \binom{k}{m} (k - 2m) \exp(j(u_ox + v_oy)(k - 2m)) \\ jv_oy \sum_{k=0}^{\alpha} \binom{\alpha}{k} 0.5^k \sum_{m=0}^{k} (2j)^{-k} (-1)^m \binom{k}{m} (k - 2m) \exp(j(u_ox + v_oy)(k - 2m)) \end{array} \right] \]

The equations above can be modified to include real \( \alpha \), using the binomial theorem

\[ (a + b)^\alpha = \sum_{m=0}^{\infty} \binom{\alpha}{m} a^m b^{\alpha - m} \]

and the following definition:

\[ \binom{\alpha}{m} = \frac{1}{m!} \prod_{n=0}^{m-1} (\alpha - n) \]

which holds for any real \( \alpha \), and any \( a \) and \( b \) such that

\[ |a/b| < 1 \]

Taking \( a=0.5\sin(u_ox+v_oy) \) and \( b=0.5 \) in (3-31), we can see that (3-33) does not hold for
\[ x = \pm (2k_1 - 1) \frac{\pi}{2} \frac{1}{u_o} \]
\[ y = \pm (2k_2 - 1) \frac{\pi}{2} \frac{1}{v_o} \]
\[ k_1, k_2 \in \{ 1, 2, 3... \} \]

Corresponding to maxima of \( s(x, y) \). It means that the sum (3.31) does not converge. Consequently, this representation cannot be used in order to infer the properties of the Fourier spectrum. Instead, below we present results of the Discrete Fourier Transform (DFT) for different values of \( \alpha \).

The equations (3.28) and (3.30) both easily translate into the Fourier domain. Increased pressure causes the sine wave to distort—the ridges become wider and flat, the valleys become narrower. In the spectral domain, it introduces higher order harmonics of the ridge-valley frequency. This can be readily verified by evaluation of the equations above for selected \( \alpha \). The results for the case when \( \alpha \) in (3.11) is real, can be shown numerically; examples for \( \alpha = 0.5 \) and \( \alpha = 3 \) are shown in Figure 3.18 below. Similar behavior has been observed for natural fingerprint impressions reproduced by live sensors [Jiang, 2000]. Higher order harmonics appear in general whenever the sinusoid is passed through a nonlinear operator.

It may be useful to investigate other models of nonlinearities that can model the effect of pressure, for example, sigmoid functions.

Figure 3.18. The effect of different values of \( \alpha \), on Fourier spectrum; wider ridges a) and narrower ridges b) correspond to different degrees of concentration of power spectrum around 0.
Another defect model involves shift of the mean and reduction of the gray scale range. This is rather easy to illustrate for constant $c_{cr}$ and $d_{cr}$. The DC component is controlled by $d_{cr}$, and $c_{cr}$ reduces the amplitude of the signal. The contrast reduction by itself is not degrading to the image quality and can be easily reversed by normalization. However, in the presence of noise reduced contrast is typically associated with reduced signal-to-noise ratio. The case of random noise (zero mean IID additive Gaussian noise) is illustrated in Chapter 6.

Intuitively, quality defects influence all of the processing stages of a pattern recognition system. The nature of this process has not been studied in the literature. While certain properties have been presented above, there are a number of processing techniques developed specifically for fingerprint recognition systems (presented in the next chapter), and it would be beneficial to know how they are influenced by different types of defects. However, because of the significant number of them this work focuses on orientation estimation methods. Orientation features are useful in segmentation, classification and enhancement of fingerprint images. Because of their importance, it is valuable to explore the relationship between the defects, modeled using tools presented in this chapter, and their effect on orientation estimation. It can be expected that a systematic approach can lead to improved designs of recognition systems. In particular, defects present in the ridge pattern introduce uncertainty to the orientation estimates. The use of a quantitative measure of uncertainty could improve overall classification performance.
CHAPTER 4

FINGERPRINT RECOGNITION SYSTEM BUILDING BLOCKS

This chapter contains selected techniques for pattern processing and classification including modifications and new approaches developed in the course of this work. In Section 4.1 certain aspects of collecting fingerprint images are presented with focus on quality assessment of sensors based on work conducted at NIST and MITRE Tech. In Section 4.2 an algorithm for segmentation of fingerprint images is introduced. Section 4.3 contains a presentation of two approaches to estimation of the orientation field. The algorithms use quality information in the form of variance of local orientation estimates. Subsequently it is shown how the orientation field can be used to find higher level features: SPs and OFFCs. The SP extraction method based on the ratio of coherence to local orientation difference is introduced. It is followed by a ridge flow tracing algorithm that allows for extraction of OFFCs. In Section 4.4 two algorithms for fingerprint categorization are presented based on SPs and OFFCs. From the literature and tests conducted in the scope of this work, it has been determined that the OFFCs-based algorithm shows very low classification error rate even in the case of partial images obtained from live sensors. The modifications to the original algorithm, proposed in this work, provide a robust method for discrimination between the two loop types. The results of the classification using datasets characterized by different quality levels are presented.

4.1 Pattern Collection

Fingerprint recognition algorithms are designed according to their different application environments. Different application environments define a range of available means of image collection. Probably the most general classification divides fingerprint recognition systems into off- and on-line. In offline systems, fingerprint images are collected independently of the system and are entered to the system manually or automatically in batch processing. It is relatively easy to define compatibility requirements that allow using images from different sources. A good example of such a system is FBI's Integrated Automated Fingerprint Identification System (AFIS). Historically, the fingerprints were collected in a form of rolled ink impressions on paper cards and a large number of them have been collected and stored in archives. The technology
available today allows for the direct acquisition of images in digital form, which in turn allows for more efficient processing. Consequently, the ink impressions are nowadays converted to digital form. In off-line systems, the requirements of processing speed and resources are not critical. However, since such systems often operate on large scale databases, the requirements on recognition accuracy are high. One of the factors that affects recognition rate is quality of the images entered into, and stored, in the system databases. If the system works in offline mode, the suitability of the input images can be decided by a human operator based on certain quality guidelines. The reliability of the match (or often a ranking of matches) reported by the system is usually verified by a professional. If a match between two low quality images is reported, it may require further examination and, as a result, it can be considered unreliable. Online systems, on the other hand, require automated mechanisms for quality evaluation. Online systems deployed at present are not implemented in large-scale systems. Examples can be found in private and public institutions and in personal electronics. In such systems, usually one type of fingerprint scanner is used. It simplifies subsequent processing and feature extraction. There are standards that define the requirements for the fingerprint scanners. The summary based on the MITRE Tech report and NIST and FBI standards is presented in the table below. The specifications are verified in a test setting against test patterns. These specifications do not require testing of the interaction between the user and the sensor. Even though quality defects can be caused by collection mechanisms, one should take into account as well the users, and the application environment. For the purpose of this study, a number of fingerprint images from FVC and NIST databases have been evaluated. The experience thus gained allowed for identification of important quality defect modes, as presented in Chapter 3. The defect models arising from the technology used in the sensing devices, such as scanners, or live sensors, is beyond the scope of this work. However, selected aspects of the performance of fingerprint scanners are presented below.

There are three types of fingerprint impressions that can be encountered: rolled paper impressions, flat impressions and latent impressions. Latent fingerprints form a separate category and are particularly challenging in the process of recognition. They consist of a residue that remains after the fingertip was in contact with a certain type of surface. In order to be used for recognition, they need to first be “lifted” and subsequently digitized. In some situations, latent impressions can cause a residual image to appear on the fingerprint scanner output (if, for instance, the scanner’s surface has not been cleaned between scans). Rolled ink impressions have a long tradition and are used in criminal records. As such, they are collected in highly controlled environments and can be expected to be of good quality, i.e. with good visibility of ridge details,
containing the complete area of the fingerprint and have a low level of elastic deformations – at least the above can be ensured with proper care. On the other hand, flat fingerprint impressions (those created as a result of a contact of a fingertip with a smooth surface) often represent a fraction of a fingerprint area and are likely to be deformed to a more significant degree than rolled impressions. For the purpose of automated fingerprint recognition, in systems such as FBI’s IAFIS, rolled ink impressions need to be digitized. Standards containing detailed requirements are available from FBI and NIST. Table 4.1 below contains sensor parameters, and has been prepared based on a publicly available technical report on the verification of fingerprint scanners [Nill, 2005]. The report contains a proposal for the methodology based on FBI and NIST standards for verification of conformance of the fingerprint and palm scanners.
Table 4.1. Parameters of fingerprint sensors.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gray level resolution</td>
<td>Number of bits used to represent gray levels</td>
<td>8 bit</td>
</tr>
<tr>
<td>Linearity</td>
<td>Transfer function between the gray level pattern and sensor output should be defined by a linear relationship</td>
<td>allowed deviation from linearity ≤7.65 levels</td>
</tr>
<tr>
<td>Spatial resolution</td>
<td>Relationship between physical and sampled resolution, expressed as a number of samples per unit length</td>
<td>500, 1000 dpi</td>
</tr>
<tr>
<td>Geometric accuracy</td>
<td>Measures how accurate geometric structures (parallel bars) present in a test image are reproduced at the scanner output. (Values in the next column are given for 500 dpi scanner.) The requirement defines 1) allowed distortion of the distance between parallel bars (shown in the next column) and 2) allowed deviation from a line within a 1.5 inch line segment. In both cases referred to above the requirements need to be fulfilled in at least 99% of measured locations.</td>
<td>Across-bar (in inches): D ≤ 0.0007, for 0.00 &lt; X ≤ 0.07 D ≤ 0.01X, for 0.07 ≤ X ≤ 1.50 Along bar (in inches): D ≤ 0.016</td>
</tr>
<tr>
<td>Spatial-frequency response</td>
<td>Expressed as a modulation transfer function (MTF). Is measured to determine how well different frequency components are reproduced on the scanner output. Minimum and maximum modulation levels are defined for different frequencies</td>
<td></td>
</tr>
<tr>
<td>Gray scale uniformity</td>
<td>Measured for 3 cases: 1) adjacent column/row, 2) pixel to pixel and 3) small area. Specifications for 1) need to be fulfilled in at least 99%, 2) 99.9% and 3) 100% of cases.</td>
<td>The threshold values are defined for low and high reflectance cases. In the case 1) of adjacent row/column uniformity, the average intensity levels between each neighboring 0.25 inch long segments can differ by at most 1 gray level for low reflectance test pattern and 2 gray levels for high reflectance test pattern.</td>
</tr>
<tr>
<td>Signal-to-noise ratio (SNR)</td>
<td>Defined as an average output signal range to the noise standard deviation. Standard deviation is computed separately for uniform black and white test images. Each of the two SNRs should be greater or equal to the threshold specified in the next column in 97% of the cases within each block.</td>
<td>≥125</td>
</tr>
</tbody>
</table>

Table 4.1 does not provide a complete description of the parameters of scanners and standards; details can be found in [Nill, 2005]. The parameters presented in the table represent nominal values that need to be satisfied for test patterns, and in practice it can be expected that the fingerprint sensor vendors ensure high performance. The geometric accuracy is measured in reference to a test pattern referred to as a target. The quantity $D$ denotes the difference between distances measured in the target and the same distance measured in the output image. $X$ denotes distances between bars, and can be viewed as a feature size. Note that different accuracy is expected for different feature sizes. Spatial frequency response is measured using modulation $m$ defined as:  

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Here $s_{\text{max}}$ denotes the peak and $s_{\text{min}}$ the valley taken as average intensities in a window. Modulation $m$ is required to be within a specified range and the minimum threshold decreases as a function of frequency. Interestingly, the report gives $m=1.05$ as maximum allowed modulation in order to limit use of preprocessing at the sensor level, such as edge enhancement, which produces high levels of modulation for certain frequency components.

From the practical point of view, fingerprint quality defects originate from interactions between sensor and user and are not solely a function of the sensor parameters. In real life applications, the output image quality can be expected to be lower. Standards defining imaging performance in the case of dry or wet prints are not available as of now; however, the issue of quality of fingerprint scans is addressed in [Nill, 2005]. The highlighted aspects of quality include: 1) the gray level range, 2) presence of artifacts and abnormalities and 3) sharpness and detail rendition. The gray level range can differ from the one specified by the nominal binary precision of the scanner. The quality specifications establish the percentage of images that are captured by the sensor that need to exceed certain numbers of intensity levels. According to [Nill, 2005], 80% of images need to have at least 200 gray levels and 99% at least 128 levels (defined for a dynamic range of 8 bits per pixel). The presence of artifacts and defects is difficult to assess quantitatively. The requirement states that they should not be significant enough to impair the fingerprint comparison. Similarly, for the sharpness and detail rendition category, the requirement is to provide enough details to allow for fingerprint comparison. The quality of a sensor is evaluated based on a set of master images. The main difficulty in quality evaluation is the ability to distinguish the fingerprint quality from the image quality impairments introduced by the scanner. Suggested methodologies for quality evaluation include MITRE IQM [Nill and Bouzas, 1992] and NIST quality assessment algorithm [Tabassi et al., 2004].

The requirements for live sensors should follow these guidelines. In practice, the verification of live sensors can be conducted to a limited extent and may require different test patterns than do optical scanners used for ink impressions. Because of the multiplicity of scanners for live fingerprint sensing, the test procedures for different sensors need to be designed specifically for
the sensing technology. [Maltoni et al., 2003] provide the survey of fingerprint sensor vendors and of the sensing technologies and their classification.

The most common live sensors available today operate in the resolution of ~500 dpi. It is believed that higher resolution, greater than 1000 dpi can lead to improvement in recognition accuracy. Partly, this is due to ability of these sensors to discern fine details of ridge lines that remain undetected at lower resolutions. One of the critical parameters of the live sensors is the size of the sensing area. The size of the scanned image for fixed resolution varies for different models of scanners. For example the array of pixels in the largest image used in FVC 2004 was 640x480 pixels (500 dpi sensor). In certain applications such as portable or handheld devices it is important to minimize the physical dimensions. They can be reduced only to a limited degree beyond which the number of details decreases too much and the recognition performance decreases to an unacceptable level [Maltoni et al., 2003]. One way to keep the resolution and image size intact while decreasing the physical dimensions of the device is by using so-called sweep sensors. A sweep sensor consists of a thin strip of detectors that take snapshots of the fingerprint as it is moved over its surface. The result is the collection of slices of the fingerprint. The slices are subsequently assembled into the complete image (a process referred to as mosaicing). Clearly, there is a potential for additional quality problems as one can expect difficulties in registration of the slices.

Available sensors use the optical, solid state and ultrasound technologies. Within each category more types can be distinguished, as shown in the table below [Maltoni et al., 2003]:

<table>
<thead>
<tr>
<th>Sensing technology</th>
<th>Sub category/ sensing method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optical</td>
<td>Frustrated total internal reflection</td>
</tr>
<tr>
<td></td>
<td>Frustrated total internal reflection with sheet prism</td>
</tr>
<tr>
<td></td>
<td>Optical fibers</td>
</tr>
<tr>
<td></td>
<td>Electro-optical</td>
</tr>
<tr>
<td>Solid state</td>
<td>Capacitive (static)</td>
</tr>
<tr>
<td></td>
<td>Thermal</td>
</tr>
<tr>
<td></td>
<td>Capacitive (dynamic)</td>
</tr>
<tr>
<td></td>
<td>Piezoelectric</td>
</tr>
<tr>
<td>Ultrasound</td>
<td></td>
</tr>
</tbody>
</table>

The collection mechanisms determine the properties of the fingerprint impression and of the background. These properties need to be known in order to design efficient processing algorithms. Selected techniques are presented in the reminder of this chapter.
4.2 Preprocessing

Preprocessing operations in fingerprint recognition systems involve segmentation and enhancement. In image processing, segmentation denotes a process in which the portrayed scene is partitioned into regions. Each region is considered to be uniform according to some criterion, which can be based, for example, on texel\(^1\) orientation, local spatial-frequency or color. In the case of fingerprints, segmentation is relatively simple. Usually scanned fingerprint images are clearly distinguishable from background, with the exception of low quality images with low contrast and high levels of noise. The techniques that can be used to detect and separate the fingerprint silhouette from background are presented in Section 4.2.1. Below, the existing techniques, previously reported in the literature, are presented briefly and modifications are subsequently proposed. The modifications include new segmentation measures and an adaptive algorithm that is independent of the sensing technology. Segmentation is instrumental in automated scene analysis. Instead of processing the entire image, one can restrict focus only to the regions/objects of interest. This task is accomplished by first computing local descriptors based on pixel intensities and then by applying an algorithm that groups together regions that share common properties. In the case of fingerprint images, the segmentation descriptors may include directional information such as histogram of orientations, energy in the direction perpendicular to ridge flow or isotropic information such as mean of intensities or histogram range [Maltoni et al., 2003]. Depending on the pattern type, it may be useful to use a vector of descriptors. The universal approaches to segmentation involve Markov Chain Monte Carlo simulations [Tu and Zhu, 2002] and biologically inspired ones [Grigorescu et al., 2002].

Image enhancement, in general, can refer to any procedure designed to improve the appearance of the image by improving the visibility of details. A good example of enhancement is image denoising. The fingerprints consist of a specific type of pattern that is characterized by spatial periodicity. This property can be used in fingerprint enhancement. General purpose image enhancement techniques include statistical approaches within a Markov Random Field (MRF) formulation. Briefly, a potential function is defined over a neighborhood for each pixel in the image. This function is then minimized, which leads to an image containing a lower level of noise. The potential function typically involves a regularization term as a consequence of the assumption that local variation in pixel intensities is low in noise-free images. An interesting and

\(^1\) texel – "similar textural elements that are replicated over a region of the image", [Davies, 2005]
relatively new technique for universal denoising involves learning the signal structure from the noisy signal itself [Weissman et al., 2005], [Yu and Verdu, 2006].

In the case of fingerprints specifically, the general purpose techniques are often used in early stages of preprocessing: normalization, contrast manipulation, Wiener filtering [Maltoni et al., 2003]. The most successful techniques rely on contextual filtering. Contextual filters are matched to local frequency and orientation of the ridge pattern, and are robust to noise and oriented artifacts. Examples include Gabor filters. All contextual filters necessarily lead to suppression of ridge structure in highly curved regions, at SPs and minutiae. This effect can be limited when the spatial frequency range of the filter is not too narrow. Another problem is related to the low quality. If enhancement is to be performed in a low quality region, where local orientation and frequency estimates are not reliable, the filtering may produce undesired results. The Section 4.2.2 presents enhancement techniques and proposes modifications that allow for a more robust enhancement. More importantly, the enhanced images do not require thresholding and output images are gray-scale images corresponding to the original image with a decreased amount of noise (here noise is understood as local variations of intensities along ridges and the usual additive Gaussian noise). In very low quality images, the larger context can be used and the enhancement effect can be increased for better noise suppression. This comes at a tradeoff between enhancement and suppression of the ridge structure, and introduction of spurious ridges.

4.2.1 Segmentation

Before introducing the segmentation methodology, it is helpful to introduce notation used throughout this chapter. Let I denote an intensity image and x and y the integer indices of pixels within the image. The indexing convention follows matrix form.

\[ I(x, y) \in \{0,1,2...255\} \]
\[ x, y \in Z \]

In our case, the set Z is often restricted to the positive integers \(Z^+\). Let us further denote by \(\Lambda\) a grid of points over the image I. In the case of a regular grid

\[ \Lambda = \{\lambda = (x_\lambda, y_\lambda): x_\lambda = k \cdot c, \quad y_\lambda = l \cdot c \quad \text{for} \quad k \leq K, \quad l \leq L \quad \text{and} \quad k,l \in Z^+\} \]
where $c$ is a constant denoting the distance between two neighboring grid points, and $k$ and $l$ are positive integers, and $K$ and $L$ represent the number of rows and columns, respectively.

Processing techniques presented in this work rely on locally computed quantities. For this purpose we define a set of blocks $B_\lambda$ centered on the grid points $\lambda = (x_\lambda, y_\lambda)$.

$$B_\lambda = \left\{ (x, y) : |x - x_\lambda| \leq \left\lfloor \frac{c_b}{2} \right\rfloor, \quad |y - y_\lambda| \leq \left\lfloor \frac{c_b}{2} \right\rfloor \right\}$$  \hspace{1cm} (4-4)

$B_\lambda$ is a square block with side length $c_b$, where $c_b = c$ for non-overlapping blocks. The $c_b$ is assumed to be an odd integer. Consequently we will denote by subscript $\lambda$ any quantity computed using the intensities in block $B_\lambda$, i.e. $m_\lambda = f(I(B_\lambda))$.

In the case of fingerprint images, the segmentation approach is based on the observation that a fingerprint is characterized by a periodic pattern of ridges and valleys (Figure 4.1). A number of statistics of the intensity image can be computed block-wise and their values can be used to guide the segmentation. Here we will refer to them as the segmentation descriptors. The block can be of arbitrary size. However, in this work it has been determined that the best results are obtained for block sizes between 1 and 2 ridge–valley periods. For increased accuracy of segmentation, one can introduce overlap between blocks and compute segmentation descriptors at every pixel.
Because of the quasi-periodic nature of fingerprints, we expect high variability of the intensities in the area of fingerprint. Examples of statistics calculated for blocks $B_x$ are presented in Figure 4.2. The mathematical description follows further in this section.

It has been shown that segmentation can be improved when more than one segmentation descriptor is used [Maltoni et al., 2003]. Coherence of gradient vectors, and the mean and variance of intensities have been used previously in [Bazen and Gerez, 2001], where the three descriptors are used to train a linear classifier. The training needs to be performed individually for different sensors. The classifier assigns each pixel to either foreground or background.
Subsequently, post processing that involves morphological operators is used to regularize the fingerprint silhouette and to remove holes and isolated regions. In this work a similar approach is taken with the following differences. Instead of training a linear classifier for each sensor, the training is done for each individual impression. The linear classifier is not obtained using a Least Mean Squares metric (MLS), but instead probabilities are estimated from histograms and the classification rule is based on the likelihoods and a specifically designed prior. Finally, our algorithm works block-wise rather than pixel-wise, which increases the processing speed. It is understandable that using block-wise rather than pixel-wise processing leads to rather crude approximation of the fingerprint contour, but it has been shown in fingerprint classification tests to perform satisfactorily. Better accuracy in the region of a boundary can be obtained by pixel-wise processing, which for efficiency can be restricted only to the pixels near the boundary region obtained using our method.

The segmentation approach presented in this section is simplified in that it does not require additional information about fingerprint features: local frequency and orientation.

Plot a) in Figure 4.2 shows an intensity profile along the middle row of the image a) in Figure 4.1. Plot b) shows four measures computed from pixel intensities over a rectangular grid of blocks (c=19 pixels with 11 pixel overlap (c=8)): mean - m, standard deviation - σ, and \(d_1\) (4-10) and \(d_2\). The block size is chosen to encompass a few ridges and valleys. Of particular interest is the \(d_1\) as it shows robustness with respect to variations in the contrast profile and average intensity level. It also tapers off rapidly on the fingerprint boundaries and separates background from foreground effectively.

In order to gain additional insight into the properties of the statistics mentioned above, it may be useful to examine the histogram of intensities. An image containing clearly separated ridges and valleys has two peaks or at least one can expect intensity levels to be spread over a wide range of values. On the contrary, background or areas with impaired quality where ridges are exaggerated and merged together have a histogram concentrated in a narrower range of intensities.
Figure 4.3. Magnified fragment of the fingerprint from Figure 4.1 a) and histogram of intensities b).

Figure 4.3 shows a block of a fingerprint image where ridges are not clearly distinguished — it is reflected on the histogram of intensities. The peak is in the lower range of intensities, thus reflecting the dominant proportion of ridges and the distribution is skewed towards low intensity levels.

Figure 4.4. Magnified fragment of the fingerprint from Figure 4.1 a) and histogram of intensities b).

Figure 4.4 depicts a block with slightly more even widths of ridges and valleys — the two clusters in the intensity histogram begin to emerge.

Ideally, one would estimate distributions of pixel values for ridges and valleys as in [Chen et al., 2004]. However such an approach requires orientation estimation and the knowledge of location of ridges and valleys \textit{a priori}. $d_1$ measures distance between two distributions and can serve as a simplified substitute. First, a mean value estimate for a block of interest is calculated, and subsequently pixels of intensities greater than the mean are labeled as "valleys" and lesser ones as
“ridges” (4-5) - (4-7). Mean and variance is computed separately for “ridges” and “valleys” (4-8) - (4-9) and for the two truncated distributions $d_i'$ is calculated (Figure 4.2) according to (4-10).

The mean intensity in block $B_i$ is defined as

$$m_i = \frac{1}{|B_i|} \sum_{(x,y) \in B_i} I(x,y) \quad (4-5)$$

Using $m_i$ as a threshold, the pixels are divided into two clusters $B_{r,i}$ and $B_{v,i}$

$$B_{r,i} = \{(x,y) : I(x,y) < m_i, (x,y) \in B_i \} \quad (4-6)$$

and $B_{v,i}$

$$B_{v,i} = \{(x,y) : I(x,y) \geq m_i, (x,y) \in B_i \} \quad (4-7)$$

For each of the clusters, means and variances are computed

$$m_{r,i} = \frac{1}{|B_{r,i}|} \sum_{(x,y) \in B_{r,i}} I(x,y), \quad m_{v,i} = \frac{1}{|B_{v,i}|} \sum_{(x,y) \in B_{v,i}} I(x,y) \quad (4-8)$$

$$\sigma_{r,i}^2 = \frac{1}{|B_{r,i}|} \sum_{(x,y) \in B_{r,i}} (I(x,y) - m_{r,i})^2, \quad \sigma_{v,i}^2 = \frac{1}{|B_{v,i}|} \sum_{(x,y) \in B_{v,i}} (I(x,y) - m_{v,i})^2 \quad (4-9)$$

These statistics are used in the following formula that measures the distance between two distributions

$$d_i' = \frac{|m_{r,i} - m_{v,i}|}{\sqrt{0.5(\sigma_{r,i}^2 + \sigma_{v,i}^2)}} \quad (4-10)$$

Alternatively, assuming there are two distinctive clusters in the histogram, the k-means algorithm [Bishop, 2006] can be used to identify them and (4-10) is subsequently applied to the two
clusters ($d_2$ in Figure 4.2). This approach, referred to as $d_2$ in this work, has been rejected as it resulted in very low values obtained from (4-10) and did not provide as good separation of foreground from background as $d_1$ does (refer to Figure 4.5). The histograms of mean, standard deviation, $d_1$ and $d_2$ below have been prepared from block-wise estimates on the image from Figure 4.1. They show clearly how regions associated with background and foreground form distinct clusters. In Figure 4.5 a), b), c) and d) the thin and spiky cluster corresponds to the background and the wide one to the fingerprint area; $d_1$ provides the best separation of the clusters.

Figure 4.5. Histograms of measures calculated for the image shown in Figure 4.1.

The images in Figure 4.6 show the spatial distribution of the statistics calculated for the image in Figure 4.1. It can be seen that those simple measures can be rather efficient in fingerprint silhouette detection.
Figure 4.6. Segmentation measures, where low intensity denotes low value and high intensity high value for a) mean, b) standard deviation, c) $d'_1$ and d) $d'_2$.

Given an image feature computed as shown above, the issue of how to choose a threshold arises. For a particular input sensor, one may determine a fixed threshold that provides satisfactory results. However, this approach is likely to yield undesired results when applied to images captured using different sensors, or even for images of different quality from the same sensor.
Figure 4.7. Segmentation results obtained when using fixed threshold: a) mean (m<0.44), b) standard deviation (σ>0.5), c) \(d_1^j > 0.5\), d) \(d_2^j > 0.5\). Image f05 from NIST 4.

A method for an adaptive threshold selection is presented below. As shown above, \(d_1^j\) responds differently to foreground and background regions and forms two distinct clusters in a histogram calculated over the image containing fingerprint and background. These clusters can be found using unsupervised learning methods. Here, for the sake of illustration, the k-means algorithm is used. The k-means algorithm returns two clusters: 1) \(F\) - the set of all blocks \(B_x\) corresponding to the foreground and 2) \(\bar{F}\) - the set of all blocks \(B_x\) corresponding to the background. Let us denote a segmentation descriptor by \(c_j\). It can assume a vector form i.e. \(c_j = (d_{1,j}, \sigma_j^2, m_j)\). The densities \(p(c_j \mid F)\) and \(p(c_j \mid \bar{F})\), estimated using Gaussian Parzen windows, correspond to foreground and background respectively, where the term foreground is used to refer to the fingerprint silhouette. It needs to be noted that in reality more than two clusters can occur. To handle such cases properly, more sophisticated techniques are necessary in order to find these clusters: affinity propagation [Frey and Dueck, 2007] or mixture model estimation [Gelman et al., 2003].

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
When applying a Bayesian framework, one is required to specify a prior probability distribution. It has been observed (Figure 4.5) that within the foreground area, local variations in calculated statistics are larger than in the case of the background locations. The amount of variation can be measured efficiently using gradient magnitude $|g|$. The prior can be constructed as a function of the gradient magnitude, so that for the background, where $|g|$ tends to assume lower values, the prior probability is high. The gradient components in the x and y directions $[g_x, g_y]^T$ are estimated using the Sobel operator [Davies, 2005].

$$|g| = \sqrt{g_x^2 + g_y^2} \tag{4-11}$$

Let the operator applied locally be denoted as a pair $h_x(m,n)$ and $h_y(m,n)$ corresponding to kernels acting along the X and Y axes respectively. Assuming the gradient operator is centered on $\lambda$ (using grid $\Lambda$, where $c=1$, i.e. the grid coincides with all image points).

$$g_x(\lambda) = \sum_{m=-l}^{l} \sum_{n=-l}^{l} f(\lambda + (m,n)) \cdot h_x(m + 2, n + 2) \tag{4-12}$$

$$g_y(\lambda) = \sum_{m=-l}^{l} \sum_{n=-l}^{l} f(\lambda + (m,n)) \cdot h_y(m + 2, n + 2) \tag{4-13}$$
The boundary effect is handled by simply restricting $\lambda$ so that the summations in the equations (4-12) and (4-13) are contained entirely within the domain of $I$.

In addition, in order to reflect the uncertainty due to the number of points used to estimate likelihoods, the argument in the prior is a weighted function of the number of points used in likelihood estimation.

$$\eta_f = \left| \frac{F}{\Lambda} \right|$$

$$|\Lambda| = |F| + |\bar{F}|$$

where $0 < \eta_f < 1$ denotes the multiplicative constant, and the prior for the background and the foreground assume the following form:

$$p(F) = \exp \left( - \left( \frac{|F|}{\sigma} \right)^2 \right)$$

$$p(F) = 1 - p(\bar{F})$$

where $\sigma$ is selected in an exploratory way. In this work, $\sigma=1$ was used with good results. Setting $\sigma$ high leads to a classifier that is more likely to classify the fingerprint region as background, whereas selecting small $\sigma$ results in greater probability of accepting background as a fingerprint region. Given the prior of (4-16) and the normalization constant

$$p(e_i) = p(e_i | F) \cdot p(F) + p(e_i | \bar{F}) \cdot p(\bar{F})$$

one can make a decision to assign a given location to foreground or background by comparing the obtained posterior probabilities. One can avoid evaluating (4-17) by performing a likelihood ratio test at each location. Note that the threshold is not only adaptively selected for the image based on likelihoods estimated for the clusters obtained from k-means, but it also depends on the location within the image via means of the gradient magnitude. We can use the Bayes theorem (4-18) to estimate the posterior probability,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
\[ p(F|c_x) = \frac{p(c_x|F) \cdot p(F)}{p(c_x)} \quad (4-18) \]

The test is easy to perform: at each location \( \lambda \), if the relation \( p(F|c_x) > p(F|c^\lambda) \) holds, or equivalently \( p(c_x|F) \cdot p(F) > p(c_x|F^\lambda) \cdot p(F^\lambda) \), then a given block \( B_x \) is assigned to the fingerprint area. It can be referred to as maximum a posteriori (MAP) segmentation.

Figure 4.9 shows the outcome of the MAP segmentation performed on a few sample images. The measures used are \( d_i \) and the standard deviation of intensities. The standard deviation used in the prior is \( \sigma = 1 \), and the block size and overlap between neighboring blocks have been chosen arbitrarily to encompass two ridge-valley cycles.

![Figure 4.9](image)

Figure 4.9. MAP segmentation based on standard deviation a,c), and based on \( d_i \) b,d). Image f05 from NIST 4 a,b), and image 110_7 from FVC2000 DB3b c,d).
Both $d'_i$ and standard deviation yield similar results in these cases. However, the greater separability of clusters (Figure 4.5) suggests that $d'_i$ is the more appropriate choice for the segmentation task. It should be noted that some authors use morphological operations to remove isolated background blocks within the foreground and to smooth the silhouette outline. Following this notion, more advanced techniques can be applied that use dynamic contours. For example, a cost function that promotes regular silhouette outlines and penalizes isolated regions could be defined.

As an alternative to $d'$, one can use information about spatial relationships between neighboring pixel intensities. The image is divided into pairs of pixels, referred to as cliques, which can be immediate neighbors or can be separated by a certain distance (in order to reduce correlation). The cliques are formed in both horizontal and vertical directions. The set of cliques comprises disjoint pairs and each such pair can be represented as a point on a 2-D plane. This set can subsequently be used in non-parametric probability density estimation.

It can be observed that foreground and background blocks exhibit different properties. The areas containing noise produce a probability distribution concentrated along the diagonal, while areas containing edge information contain points far from the diagonal. These properties allow distinguishing foreground (oriented edges) from background (primarily noise), as shown in Figure 4.10. A probability distribution can be estimated using Parzen windows and, as has been verified through tests conducted within the scope of this thesis, the entropy of this distribution provides sufficient discriminating capability. Given the values of entropy over the grid of all locations within the image the threshold finding techniques similar to the ones described above can be used. Some examples are presented below.
In Figure 4.11, two rather difficult cases are presented. The difficulty stems from the presence of directional structures outside the fingerprint in image a) and the combination of low signal-to-noise level within the fingerprint and the presence of directional structures outside of the fingerprint silhouette. The weakness of local descriptors used in segmentation is evident in Figure 4.12. Oriented portions of background tend to be misclassified as foreground and low noise areas of the fingerprint are left out. It suggests a need for more robust techniques.

Figure 4.11. Examples of images, a) FVC 2000 DB 3 a, and b) FVC 2000 DB 3 b.
Figure 4.12. Comparison of the outcomes of segmentation using d-prime 1 a,c) and 2-cliques b,d).

It can be seen that there are some advantages from using the cliques rather than d'. The areas of segmented fingerprint images (compare Figure 4.12 with Figure 4.11) are better represented in the cases b) and d). In all cases, some post-processing of the silhouette was required in which morphological operators have been used to fill the holes in the segmented area. Based on the experiments, however, it cannot be concluded that one of the measures is better than the other. It can be beneficial to include global criteria that enforce smoothness of the contour as in [Tu and Zhu, 2002].

As an alternative to the approach in which clustering and probability density estimation is performed on-line for each input image, the training can be done once off-line based on a large set of images. Such an approach limits time consuming clustering and likelihood estimation, and replaces them by a look-up operation. With such a look-up approach, possibly more segmentation
descriptors can be used without significantly increasing the processing time, as it was done in [Bazen and Gerez, 2001]. However, the look-up approach lacks flexibility and universality with respect to image source. It can also be concluded that a hard decision regarding segmentation may not be desired at the early processing stage. More features are typically extracted at the later stages of pattern recognition and can be used to refine the first crude estimate of the silhouette shape. The uncertainty of extracted features both local and global could be helpful.

4.2.2 Enhancement

In the defect model proposed in Chapter 3, the defects have a form of a superposition of Gaussians. It implies that they have a low pass character with the spectral extent determined by $\sigma$. Unfortunately, bandpass filtering of the signal is not sufficient to separate $S(u,v)$ from $M(u,v)$ due to the term resulting from convolution in the frequency domain. However, one can use the knowledge about the frequency and the local orientation of ridges in order to “enhance” the visibility of ridges, as is the case in contextual filtering. This technique is based on a bank of matched filters that are tuned to local frequency and orientation [Maltoni et al. 2003]. Care needs to be taken in order to avoid introduction of spurious details. In particular, the filters tuned to the orientation that does not correspond to local orientation will suppress the true orientation. There is also a known tradeoff between spatial and spectral width of the filter pass band.

An interesting method that does not require prior knowledge of the local orientation and frequency of ridges is presented below. The image is enhanced in the Fourier domain using a matched filter. The filter, however, corresponds to the signal’s magnitude raised to the $k^{th}$ power. Figure 4.13 below shows a method following this notion; the image has been enhanced using a filter constructed based on the signal itself (center) and using a modified filter constructed based on a bandpass filtered version of the signal (rightmost column). The original images are presented in the leftmost column in Figure 4.13.
Figure 4.13. The results of image enhancement using the technique presented in [Maltoni et. al. 2003] (the second column) and the modified approach in which a bandpass filtered signal was used to construct an enhancement filter (last column). First row – ridges with breaks, second row – a good quality image, next two rows – subimages taken from a low quality fingerprint image and the bottom row – a synthetic image.
The equations (4-19) through (4-21) below describe the enhancement process.

\[
\begin{align*}
\text{s}_{\text{enh}}(x, y) & \leftrightarrow S'(u, v)^{FT} \cdot |S'(u, v)|^k \\
\text{s}_{\text{enh}}(x, y) & \leftrightarrow S'(u, v)^{FT} \cdot |S'_{bp}(u, v)|^k \\
S_{bp}^*(u, v) & = S^*(u, v) \cdot H_{bp}(u, v)
\end{align*}
\] (4-19) (4-20) (4-21)

Here, \(s'(x, y)\) is the observed image, \((x, y) \in B_x\), \(S'(u, v)\) is its Fourier Transform, \(s_{\text{enh}}(x, y)\) is the enhanced image, \(k\) is the coefficient, usually not greater than 1.4. In the example in Figure 4.13, \(k=1\). \(H_{bp}(u,v)\) is a Butterworth bandpass filter (Figure 4.14). As can be seen, using a bandpass filtered version of the signal’s magnitude results in slightly better noise suppression and also removes slow variations in intensity. The coefficient \(k\) controls the amount of enhancement and the amount by which the image structure is altered; larger values lead to better suppression of noise and other artifacts, but setting it too high leads to an increased number of filtering artifacts.

It can be seen that the filter based on signal magnitude raised to a power tends to shape the image according to the strongest frequency component in the image. As a result, fine details such as minutiae or ridges near SPs can be inadvertently removed. In the middle column this effect is not as significant as in the right column, but can nevertheless be found in both cases in the area of the ridge bifurcation (row two and three). Consequently, the bandwidth of the filter should be chosen as a tradeoff. A narrow band assures robustness to noise, but it tends to distort an image in the regions of high curvature. A wider band, on the other hand, tends to preserve fine details at the cost of weaker signal improvement.
4.3 Feature Extraction

The subject of this section is to present methods for extracting features from the ridge orientation field that can be used in the fingerprint categorization task. In a given classification task, the problem that arises is how to find the set of features that provides the best possible discriminatory information and subsequently how to reliably extract the set of features from an image. Available tools include independent component analysis (ICA), PCA, Fisher's linear discriminant [Duda et al., 2001, Bishop, 2006] for feature selection and dimensionality reduction. Numerous statistical tools, such as finding mixture components, can be applied [Gelman et al., 2003], as well as clustering methods, such as the k-means method [Duda et al., 2001], or the affinity propagation algorithm [Frey and Dueck, 2007]. These are general purpose tools. Sometimes it is more convenient to use features used by human examiners. Such heuristics can be successfully used in automated fingerprint categorization, where the useful features can be computed from the orientation field, which in turn can be obtained from local ridge line orientation estimates.

Below we present methods that allow for detection of SPs and for extraction of OFFCs [Dass and Jain, 2004]. A combination of these two types of features allows for robust categorization of fingerprints. For the illustration of the quality impact on the classification, only OFFCs are further used. SPs as discriminatory features can be unreliable in low quality images. They allow for 3-class categorization; it is not possible to distinguish tented arches from loops using solely SPs,
whereas OFFCs are more robust. They contain sufficient information to distinguish left loops from right loops, thereby allowing for 4-class categorization. If one includes SPs and the approach presented in [Karu and Jain, 1996] for discriminating between loops and tented arches, one can obtain a 5-class categorization scheme. Additionally, since OFFCs rely on the global orientation pattern, they are more robust to local orientation field distortions. These can still be harmful to a degree depending on a ridge-line tracing method. The classification method used as a case study in this thesis is based on features of OFFCs. The tracing method used in this work is rather sensitive to local field distortions. It is therefore important to obtain a robust estimate of the orientation field first. The two methods presented below use quality information and properties of the fingerprint ridge line patterns.

The ridge orientation field of a fingerprint can be assumed to be locally smooth with the exception of SPs and defects. Such assumptions lead to good orientation field estimation results [Perona, 1998, Wang and Wang 2004, Cappelli et al., 1999]. In fact, a straightforward lowpass filtering [Bazen and Gerez, 2002, Chen et al., 2004, Hong et al., 1998, Ando, 2000, Jain et al., 2000] of gradients or orientation field leads to robust orientation estimates. It is our desire to find methods of utilizing the dependence among neighboring blocks in a more structured manner than the lowpass filtering. Of particular interest are probabilistic models of the 2-D field. If defects are present, and they are significant enough to distort the orientation field, one remedy is to increase the size of the analysis window [Jain et al., 1997] or adapt scale, shape and orientation properties of the operator (cf. the scale-space approach presented in [Almansa and Lindeberg, 2000]). Two new methods for smooth orientation field estimation are presented below. Our first algorithm performs the orientation estimation in an iterative manner. [Karu and Jain, 1996] proposed iterative averaging in order to regularize the orientation field. The averaging is stopped when number of SPs detected in the image is reasonable. The idea presented here is to smooth the image not by iterative averaging, but by increasing the size of the analysis window in subsequent iterations. The idea is to use larger neighborhoods to suppress noisy artifacts. This method leads to natural selection of the block size guided by improvement in smoothness of the orientation field rather than by a number of detected SPs that is not known a priori. There are two stages of the algorithm. In the first stage the orientation estimate is calculated while the block size $c_b$ is gradually increased (without introducing an overlap between blocks). A cost function $\theta_D$ (4-22) is calculated using all locations in the image at each iteration. The block size is increased until a threshold is met or if there is no significant decrease in $\theta_D$. The convergence of the method to a minimum is a result of a compromise between increasing smoothness (achieved by means of
averaging of defects and noise) and increasing local orientation differences in the neighborhood as neighboring block centers move apart (Figure 4.16).

\[ \theta_D = \frac{1}{|\Lambda|} \sum_{\lambda} \theta_{d,\lambda} \quad (4-22) \]

Subscript \( \lambda \) denotes the neighborhood center and \( |\Lambda| \) is the cardinality of the set of all blocks centered on \( \lambda \in \Lambda \).

\( \theta_{d,\lambda} \quad (4-23) \) is a mean orientation difference in the neighborhood of \( B_\lambda \). In our algorithm, local orientation difference is a measure of ridge curvature. This measure was previously applied in [Cappelli et al., 1999] in a different context, namely to express a difference between average orientations of two fingerprint regions. \( N_\lambda \) denotes the neighborhood of \( \lambda \) and the index \( \lambda' \) denotes locations that belong to \( N_\lambda \). The angle \( \theta_\lambda \), here measured in degrees, is the local ridge orientation at location \( \lambda \).

\[ \theta_{d,\lambda} = \frac{1}{|N_\lambda|} \sum_{\lambda \in N_\lambda} |\theta_{d,\lambda'}| \quad (4-23) \]

\[ N_\lambda = \{ \lambda': |\lambda' - \lambda| = 1 \} \quad |N_\lambda| = 8 \quad (4-24) \]

\[ \theta_{d,\lambda'} = \begin{cases} \theta_\lambda - \theta_{\lambda'}, & -90 \leq \theta_\lambda - \theta_{\lambda'} \leq 90 \\ \theta_\lambda - \theta_{\lambda'} - 180, & \theta_\lambda - \theta_{\lambda'} \geq 90 \\ \theta_\lambda - \theta_{\lambda'} + 180, & \theta_\lambda - \theta_{\lambda'} \leq 90 \end{cases} \quad (4-25) \]

Other measures such as coherence of orientations [Bazen and Gerez 2002] can be considered as alternatives. Figure 4.16 below shows how the cost function \( \theta_D \) defined in (4-22) changes with increasing block size \( c_b \). For illustration we selected 4 images from FVC2004 Db1b as presented in Figure 4.15.
Figure 4.15. Fingerprint images (FVC 2004 Database 1 b); different impressions of the same fingerprint ordered in decreasing perceptual quality.

The curves in Figure 4.16 are labeled to match fingerprint images from Figure 4.15.

Figure 4.16. Cost function $\theta_D$ as a function of block size.

Note that the perceived quality ordering corresponds to the position of the curve; for the top quality image shown in Figure 4.15 a), the cost function takes on the lowest values. This ordering property suggests that a quality measure can be defined as a function of: 1) the block size for which the minimum of $\theta_D$ is observed, and 2) the relative position of the $\theta_D$ curve for different impressions of the same fingerprint. The behavior of $\theta_D$ in Figure 4.16 can be explained as a result of two phenomena: 1) the increase in block size reduces the effect of defects (small irregularities in ridge orientation) and thus increases the smoothness of the orientation field ($\theta_D$ is limited from below by the structure of the fingerprint impression); 2) the block centers are moving farther apart as block size increases, thus resulting in larger differences of orientations of neighboring blocks and hence decreasing the smoothness. The stop condition is defined in the following way: 1) the cost function has value below threshold $t_1$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
or 2) the improvement in smoothness is less than a threshold $t_2$ (the first derivative is sufficiently small)

$$\theta_{D,(n-1)} - \theta_{D,(n)} < t_2$$

The subscript $n$ in parentheses denotes the iteration number. Figure 4.17 shows histograms of block sizes calculated for two datasets from FVC [Maio et al., 2002].

![Figure 4.17. Histograms of block sizes obtained in the first stage, without overlap between blocks (o – FVC2000 DB3b, x – FVC2004 DB1b).](image)

Note the peaked shape and the location of the maximum for FVC2004 DB1b (x), which is considered to contain images of better quality on average than FVC2000 DB3b. The latter was found to be the most difficult in identification tests [Maio et al., 2002]. In order to further increase smoothness in the second stage of the algorithm, starting from the block size $c_0$ arrived at in the first stage, an overlap can be introduced between blocks. The overlap between blocks is increased while maintaining the block centers fixed. The iterations are continued until a stop condition (defined similarly as for the first stage) is met. From Figure 4.18 it can be seen that the orientation estimate is made more reliable in the second stage. This effect is more readily observable in low quality fingerprints.
The negative aspect of this method is a drift of SPs taking place when the block size is required to be too large. This effect was noted previously in the literature. In order to remedy it, a hierarchical approach can be used. A possible solution includes tracking maxima in the scale-space framework [Lindeberg, 1994]. Having arrived at the smooth orientation field estimate, the grid of points can be made denser if needed. It should be emphasized that our method does not require detection of SPs as in [Karu and Jain, 1996]. In addition, in comparison to [Karu and Jain, 1996], the orientation field does not undergo multiple iterations of smoothing, but instead the estimates are found using larger windows which reflects the image structure better. The advantage of the iterative smoothing is that the orientation information is propagated over larger neighborhoods. During this process the consistent neighboring orientations remain virtually unchanged and the outliers are removed. Naturally this leads to significant alterations of the orientations around SPs, which is undesired.

Additionally, within our framework for smooth orientation field estimation, we have observed that a ratio of local orientation differences, $\theta_{d,\lambda}$, to coherence of gradient vectors (defined in the appendix), here denoted as $c_\alpha$, tends to exhibit strong maxima at singularities (Figure 4.19). As a singularity detector, this method is particularly robust in terms of false rejection rate.
Figure 4.19. Smooth orientation estimate; a) coherence of gradient vectors $c_\lambda$, b) orientation field $\theta_\lambda$, c) local orientation difference $\theta_{d,\lambda}$, and d) $\theta_{d,\lambda}/c_\lambda$ (for image in Figure 4.20).

It is important to stress that both $\theta_{d,\lambda}$ and $c_\lambda$ are assumed to be calculated using a block size (possibly with overlap) that provides a smooth estimate of the orientation field. Table 4.3 shows the top-ranked local maxima and the SPs correspond to the strongest ones.

Table 4.3. Local maxima of $\frac{\theta_{d,\lambda}}{c_\lambda}$

<table>
<thead>
<tr>
<th>NO.</th>
<th>ROW</th>
<th>COLUMN</th>
<th>VALUE AT MAXIMUM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>311</td>
<td>284</td>
<td>1454.56</td>
</tr>
<tr>
<td>2</td>
<td>257</td>
<td>293</td>
<td>597.97</td>
</tr>
<tr>
<td>3</td>
<td>302</td>
<td>320</td>
<td>104.21</td>
</tr>
<tr>
<td>4</td>
<td>356</td>
<td>374</td>
<td>91.22</td>
</tr>
<tr>
<td>5</td>
<td>401</td>
<td>356</td>
<td>83.19</td>
</tr>
<tr>
<td>6</td>
<td>410</td>
<td>338</td>
<td>62.63</td>
</tr>
<tr>
<td>7</td>
<td>392</td>
<td>329</td>
<td>56.23</td>
</tr>
<tr>
<td>8</td>
<td>212</td>
<td>374</td>
<td>34.95</td>
</tr>
<tr>
<td>9</td>
<td>212</td>
<td>356</td>
<td>29.97</td>
</tr>
</tbody>
</table>

Figure 4.20 shows two images of the same fingerprint impression with the dark boxes denoting local maxima of the ratio $\frac{\theta_{d,\lambda}}{c_\lambda}$ a) and, for comparison, SPs detected using a Poincaré-based method [Bazen and Gerez, 2002] b).
It is possible to use the precision of the orientation estimate to control the smoothing procedure as shown in the following presentation. The precision is understood as an inverse of variance which can serve as a measure of uncertainty of orientation estimate. The methods that can be used to calculate the variance for different orientation estimation methods are presented in the next chapter. Here, for the sake of illustration, a PCA of gradients has been used. It can be observed that if one is equally certain of all orientation estimates in the neighborhood $N_2$, then smoothing using our framework is equivalent to smoothing using a uniform window. Otherwise, orientations that have lower variance, and hence higher precision, are expected to influence the final orientation estimate more strongly. The goal was to limit the undesired effect associated with averaging of the orientation field. There, the high curvature regions corresponding to good quality are affected in the same degree as bad quality regions exhibiting high local orientation changes. This could not be fully avoided in the method presented below since the quality descriptors, expressed as variance of the orientation estimate, assume low values, at and around, SPs. Using small blocks, approximately equal to 1 pixel or less than one ridge-valley distance, is helpful in decreasing this undesired effect. The following method has been tested based on the local orientation estimation obtained using a minimum mean squared error (MMSE) fit (PCA of gradient field). It is, however, applicable to any method for orientation estimation shown in the next chapter. The estimate of orientation is assumed to follow a Gaussian distribution with known variance. We further assume a conjugate likelihood and prior [Gelman et al., 2003]. The number of iterations is constrained by the amount of change of the precision parameter (inverse of variance). Our intention was to use a criterion not related to features of the orientation field as it
is typically done. This is one of the possibilities explored so far. Others may include smoothness expressed as $\theta_0$ as discussed earlier. In choosing a value of the threshold, we follow the law of diminishing returns.

This rather simple approach to orientation field estimation is based on a Markov random field formulation (MRF). It gives good results and can be considered as an encouraging starting point. The future work should incorporate multiparameter hierarchical models.

For the purpose of illustrating how variance can be used in the smoothing of the fingerprint orientation field, we model the distribution of the orientation, $\theta_{x,(n)}$, at the center of the neighborhood, $N_x$

\[
p(\theta_{x,(n)} \mid \bar{\theta}_x) \propto p(\bar{\theta}_x \mid \theta_{x,(n)}) \cdot p(\theta_{x,(n)})
\]  (4-26)

$N_x$ denotes a neighborhood and has been previously defined. The subscript $(n)$ signifies the $n^{th}$ iteration, and $\bar{\theta}_x$ is the mean orientation over all locations in the neighborhood, including the orientation at the center point $\theta_x$, as shown below in Eq. (4-27), where the $\theta$'s under the sum are taken as the most recent values.

\[
\bar{\theta}_x = \frac{1}{|N_x|+1} \sum_{x' \in N_x \cap x} \theta_{x'}.
\]  (4-27)

Note that the orientation is circularly periodic and needs to be handled accordingly (see, for example, [Cappelli et al., 1999; Kass and Witkin, 1987]). We make the simplifying assumption of local smoothness of the orientation field, as do others [Jain et al., 1997; Dass and Jain 2004]. This assumption holds far from SPs (core and delta) where the algorithm presented in this work gives the best results. Consequently, we express the likelihood term from (4-26) as

\[
p(\bar{\theta}_x \mid \theta_{x,(n)}) \propto \exp\left(-\frac{1}{2} \frac{(\bar{\theta}_x - \theta_{x,(n)})^2}{\sigma_x^2}\right)
\]  (4-28)

The variance of the sample mean $\bar{\theta}_x$ is equal to
\[
\sigma^2 = \frac{\sum_{\lambda' \in N_\lambda \cap \lambda} \sigma^2_{\lambda'}}{(|N_\lambda| + 1)^2}
\]

where under the sum the \( \sigma \)'s from the most recent updates are used, and \( \sigma^2_{\lambda} \) corresponds to the most recent variance estimate. We assume a conjugate prior (4-30), where \( \sigma_{\lambda(n)} \) is the standard deviation of the orientation estimate at the \( n \)th iteration.

\[
p(\theta_{\lambda(n)}) \propto \exp \left( -\frac{1}{2} \frac{\left( \theta_{\lambda(n)} - \theta_{\lambda(n-1)} \right)^2}{\sigma^2_{\lambda(n-1)}} \right)
\]

Substituting (4-28) and (4-30) into (4-26) we obtain

\[
p(\theta_{\lambda(n)} | \bar{\theta}_{\lambda}) \propto \exp \left( -\frac{1}{2} \left[ \frac{(\bar{\theta}_{\lambda} - \theta_{\lambda(n)})^2 + (\bar{\theta}_{\lambda} - \theta_{\lambda(n-1)})^2}{\sigma^2_{\lambda(n)} + \sigma^2_{\lambda(n-1)}} \right] \right)
\]

and after some algebraic operations (see Chapter 6 for the solution in the case of a product of bivariate Gaussians) we arrive at the solution for \( \theta_{\lambda(n+1)} \) and variance, \( \sigma^2_{\lambda(n+1)} \), of the posterior probability of the orientation estimate at location \( \lambda \) [Gelman et al., 2003].

\[
\frac{1}{\sigma^2_{\lambda(n+1)}} = \frac{1}{\sigma^2_{\lambda(n)}} + \frac{1}{\sigma^2_{\lambda}}
\]

(4-32)

\[
\theta_{\lambda(n+1)} = \frac{1}{\sigma^2_{\lambda(n)}} \theta_{\lambda(n)} \frac{1}{\sigma^2_{\lambda}} \bar{\theta}_{\lambda}
\]

(4-33)

Equations (4-32) and (4-33) above constitute the iterative algorithm for smoothing of the orientation field using a weighted average, where the weights are precisions of the orientation.
estimates. $\bar{\theta}_\lambda$ is updated using the most recent estimates of orientations. The orientation and precision at every location in the fingerprint are updated until a stop condition is met, that is when

$$\sum_{\lambda \in \Lambda} \sigma_{\lambda,n}^2 - \sum_{\lambda \in \Lambda} \sigma_{\lambda,n-1}^2 < t \quad (4-34)$$

The chosen threshold is $t=0.1$, which in practice leads to fewer than 4-5 iterations over a reasonable quality fingerprint. Figure 4.21 shows a result of smoothing using the probabilistic model in which knowledge of variance of orientation is used to guide the smoothing. The likelihood (4-28) imposes a structure in which the orientation in the center of the neighborhood is the mean of the orientations in the neighborhood. In Figure 4.21 this leads to a shift of the location of the core. In the area of SPs the uncertainty of the orientation estimates is typically higher than in the regions far from SPs. This translates into higher variance of the orientation estimates. High variance locations are more likely to be altered in subsequent iterations.

![Figure 4.21. Smooth estimate of orientation field using MRF $\theta_0=4.73$ (block size is 17 pixels with no overlap).](image)

In the next section the categorization results are presented. Below we present the methods for SP extraction and OFFC tracing.

The most common methods of SP detection rely on orientation field estimates and use the fact that the SPs are neighborhoods in which the ridge orientation changes rapidly or in which the orientation field contains a discontinuity. The overview of the methods can be found in [Maltoni et al., 2003]. We would like to point out only the general idea for the sake of brevity. A coarse
method of SP detection can be based on a measure of curvature in a neighborhood. This has been illustrated earlier in this section in the context of the smooth orientation field estimation. It can be conducted at different scales – blockwise for coarse localization or pointwise for higher accuracy. A SP, either core or delta, can be found by calculation of irregularity [Cappelli et al., 1999] or as a coherence of local orientations, as defined in [Bazen and Gerez, 2002]; these two measures represent the same property. Using the regularity or coherence alone does not allow distinguishing between the types of SPs. This can be accomplished using templates as in [Dass, 2004], where the SPs are detected using templates simultaneously with the MRF orientation field estimation. [Bazen and Gerez, 2002] used templates to find orientations of the SPs. A coarse localization of the core can also be determined by finding a “focal point” defined by intersection of lines normal to the ridges surrounding the core [Novikov and Kot, 1998]. The commonly accepted methods rely on some form of Poincaré index-based methods. They are attractive since they allow for detection of SPs in an efficient fashion, and are able to discriminate between the two types of singularities. The Poincaré index-methods require smooth estimates of the orientation field in order to avoid spurious detections [Bazed and Gerez, 2002]. There are several variations of this method, but the principal idea remains the same: 1) choose a closed path on the $L \subset \Lambda$; 2) following the path calculate pair-wise differences of orientations $\theta_i$'s between neighboring points on L. If the sum of differences is 0 then the path does not encircle any singularity, and otherwise a proper rule is used to determine the type of SP. In the method proposed by [Bazen and Gerez, 2002], a core results in a Poincaré index of $2\pi$ and a delta has index $-2\pi$. The Poincaré index-method also allows for classification of the ridge flow type to which the detected singularity belongs [Maltoni et al., 2003]. The Poincaré index is defined as

$$PI_L = \sum_{\lambda_i \in L} \theta_{d,\lambda_{i-1},\lambda_i}$$

$$L = \{\lambda_i : i = 1,2,... | L\}$$

where $\theta_{d,\lambda_{i-1},\lambda_i}$ denotes orientation difference between consecutive points $\lambda_i$ on path L.
Figure 4.22. Examples showing singular points detected using Poincaré index-based method [Bazen and Gerez, 2002]; from left to right FVC2000 DB3b, FVC2004 DB1b, NIST Database-4.

Figure 4.22 presents examples with marked SPs (● corresponds to core and ▲ to delta singularity). The image in the center contains in addition two spurious core singularities – this is a frequent case in low quality regions that can often be observed on the boundaries of the fingerprint silhouette. Spurious detections are a direct consequence of non-smooth orientation field estimates and can be avoided by performing averaging or other regularization procedures prior to SP detection.

The following recursive rule can be used for tracing the OFFCs.

$$\lambda_i = \lambda_{i-1} + d_i \Delta_s \cdot [\cos \theta_{i-1}, \sin \theta_{i-1}]$$

(4-36)

where \(\lambda_o\) is a starting point, \(d_i=\{-1,1\}\) determines the direction in which OFFC is constructed, \(\Delta_s\) is a sampling interval and \(\theta_{i-1}\) denotes the orientation at location \(\lambda_{i-1}\). Each OFFC is traced from the starting point \(\lambda_o\) in the opposite directions as it has been proposed in [Dass and Jain 2004]. In the original approach due to [Dass and Jain 2004] the starting points are selected along vertical and horizontal lines crossing the midsection of the fingerprint silhouette. It has been slightly modified in this work. The location of \(\lambda_o\) for the subsequent OFFC is directed by local ridge orientation. A starting point for a new OFFC is located in the direction normal to the local ridge orientation at \(\lambda_o\) of the previously traced OFFC. The image is traversed from left to right and from top to bottom. If the fingerprint contains a core SP, then the subsequent starting points are moved toward it. For live impressions this is useful. [Dass and Jain 2004] tested their algorithm on a dataset of rolled impressions where the method for selection of starting points has shown excellent performance. Live impressions usually present more challenge because they often
contain a smaller portion of the fingerprint than the rolled impressions do. The proposed method ensures that the OFFC's are traced in the area that contains useful discriminatory information. Once a set of OFFCs are collected, they need to be transformed to an invariant representation, which is presented next. Each segment of the OFFC can be viewed as a unit vector:

$$e_i = \frac{1}{\Delta_s} (\lambda_i - \lambda_{i-1}). \quad (4-37)$$

Given N-1 vectors associated with each of the line segments, one can define

$$\cos(\gamma_i) = e_i \cdot e_i \quad (4-38)$$

where $\gamma_i$ corresponds to an angle between the vector describing the first segment of the OFFC and the $i^{th}$ one. The $\cos(\gamma_i)$ representation has a property that allows for distinguishing between arch, loop and whorl patterns. Two fingerprints shown below in Figure 4.23 represent a tented arch and a whorl. The OFFCs have been superimposed on the original impressions with compressed and shifted intensities for better visualization. It can be seen that for the tented arch there are OFFCs that have the properties of an arch, as desired, but some OFFCs that cross the delta point can exhibit loop properties. This creates an ambiguity and is one of the algorithm's weaknesses that preclude this approach from being immediately applicable to the case of tented arches. These problematic cases can be avoided provided that there is a way to detect OFFCs that cross singular points and discard them. One way of approaching this problem is to disallow changes in slope of neighboring segments of an OFFC greater than a certain threshold value. In this way those OFFCs that have non-smooth behavior due to a quality defect can be detected and discarded as well.
Figure 4.23. Fingerprint examples with superimposed OFFC; tented arch a), whorl b).

In the following Figure 4.24 through Figure 4.26, examples of plots of $\cos(\gamma_i)$ a), and OFFCs b) for selected OFFCs from Figure 4.23 are depicted.

Figure 4.24. Examples of $\cos(\gamma_i)$ -a), and OFFC - b) for an arch pattern.
Figure 4.25. Examples of $\cos(\gamma)$ – a) and OFFC – b) for a left loop pattern.

Figure 4.26. Examples of $\cos(\gamma)$ – a), and OFFC – b) for a whorl pattern.
4.4 Pattern Classification

The statistical approach to pattern classification is based on Bayes’ theorem about conditional probabilities. Given a pattern x from a feature space that belongs to a class ω, given class-conditional probabilities \( p(x|\omega) \) and prior probabilities on classes \( p(\omega) \), one can determine the class membership or a label of x with a probability described by the \textit{a posteriori} probability \( p(\omega|x) \). This approach ensures minimal error given that the probabilities can somehow be determined (refer to Chapter 2). It is also possible to include utilities or risk associated with taking or not taking a given action. Here we are focusing on the fingerprint categorization task, which is performed based on global characteristics of the ridge-furrow pattern captured via means of OFFCs. There are no published methods, to the author’s knowledge, that allow representing the OFFC in a form that allows posing the problem in the Bayesian framework. In Section 2.2 it has been shown that using additional information in the form of the conditional probability \( p(x|\mathcal{I}) \) can convey quality information to the classifier. In the next chapter it is shown how the quality, measured as an uncertainty of the orientation estimate, can be calculated for low level features – the local orientations. A way of representing the quality of higher level features in the form that would allow us to use it in the Bayesian classification, however, has not been determined.

In this case we are interested in finding classes: arch, left loop, right loop and whorl, to which a fingerprint belongs. For large databases the low classification error rate is of particular importance. Assignment of fingerprint patterns according to ridge-flow characteristics can be used for indexing large datasets of fingerprints, thereby resulting in an effective decrease of the number of patterns that need to be matched against, and thus reducing the access time and imposing less stringent requirements on classification error rates. [Wayman, 1999] presents a methodology for assessment of error rates for a 1-n biometric system with an indexed database.

Two approaches are presented below, one based on SPs and one based on OFFCs. Subsequently, those features are used as an input to a rule-based classifier. A number of SPs present in the fingerprint and their types are sufficient in order to determine the category of the ridge flow pattern:
### Table 4.4. Singular point-based categorization.

<table>
<thead>
<tr>
<th>No.</th>
<th>core SP #</th>
<th>delta SP #</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>0</td>
<td>0</td>
<td>Arch</td>
</tr>
<tr>
<td>2.</td>
<td>1</td>
<td>1</td>
<td>Tented arch</td>
</tr>
<tr>
<td>3.</td>
<td>1</td>
<td>1</td>
<td>Left Loop</td>
</tr>
<tr>
<td>4.</td>
<td>1</td>
<td>1</td>
<td>Right Loop</td>
</tr>
<tr>
<td>5.</td>
<td>2</td>
<td>2</td>
<td>Whorl</td>
</tr>
</tbody>
</table>

As it can be seen from Table 4.4, however, there are ambiguities; tented arch and loops are characterized by the same number and types of singularities. Consequently, one needs additional information embedded in the ridge pattern. For example, [Karu and Jain 1996] proposed to use an average of the sine of the angular differences between the slope of the straight line connecting core and delta and the local ridge orientations that the line crosses, denoted by $\gamma$. In tented arches $\gamma$ should be close to 0. The left and right loops can be distinguished by finding the relative position of the core and delta; in the left loop, ridges that encircle the core singularity are on the left side of the delta singularity, and the right loop can be viewed as a mirror image of the left loop.

The method based solely on the singularity is very sensitive to spurious SPs which occur when the orientation field is estimated over a low quality image and no corrective measures have been taken. The algorithmic approach to categorization using SPs may be represented as follows (d_cnt denotes a number of delta SPs and c_cnt – a number of core SPs):

```plaintext
if (c_cnt==0) and (d_cnt==0)
category='arch';
else if c_cnt==1
  if d_cnt>=2
    o='whorl';
  else if d_cnt==1
    if $\gamma<0.2$
      category='tented arch';
    else if (core SP positioned to the right of delta SP)
      category='right loop';
    else
      category='left loop';
  end
else
  category='loop';
end
else if c_cnt>=2
  category='whorl';
else if d_cnt==1;
  category='loop';
```

82

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
else if \( d_{cnt} \geq 2 \) 
  \[
  \text{category} = \text{'whorl'};
  \]
end

It can be seen that the algorithm accounts for the cases of missed SPs. The ordering of the rules has been determined by the way of experiments and has shown good performance. For example, it has been determined that in the images, particularly those from live sensors, the core SPs are more reliably detected than delta SPs. Consequently detection of two core SPs has been established as sufficient information for assigning ‘whorl’ category. The information regarding the presence of delta SPs is always considered after checking the occurrences of core SPs.

A more robust categorization algorithm can be developed if one takes into account the entire ridge flow pattern. Individual ridge lines can be traced and used for this purpose but as it has been shown by [Dass and Jain, 2004], one can trace the flow of the orientation field instead. The algorithm based on the latter approach has been shown to perform rather well on the NIST 4 fingerprint database of 4000 images [Dass and Jain, 2004]. It is attractive since tracing the OFFCs instead of the individual ridges is more efficient and alleviates problems related to ridge terminations and bifurcations. The OFFC features are more robust because, instead of relying on a discrete set of SPs, they involve larger areas of a fingerprint. This makes them less sensitive to local quality defects. OFFCs can be viewed as curves that are at each point tangent to the local orientation of ridges. This remains true, however, only if the sampling is chosen to be dense enough. If coarse sampling is used, then in the high curvature regions the OFFCs cross ridge lines occasionally. It can be shown that such curves have certain invariance properties to rotation and scaling, which can be achieved after translating the OFFC representation into \( \cos(\gamma_i) \) (refer to previous section).

It can be found that for an arch pattern \( \cos(\gamma_i) \) changes phase at most by \( \sim 90 \) degrees, for a loop the phase changes by 180 degrees, and for whorl there are multiple phase reversals corresponding to the number of encirclement made by OFFC around the core. An additional test is necessary in order to determine the loop type. The original approach [Dass and Jain, 2004] is to examine the sign change of the product of the coordinates of the vector \( e_r \). In this work a new approach is proposed. Suppose that for a given fingerprint two types of OFFCs have been detected: arch and loop. In order to distinguish between left and right loops it suffices to compute the distance between the ends of the arch and the loop OFFCs. Now, we can use the fact that, for the left loop, the distance between the left end of an arch OFFC and either end of the loop is less than the distance between the right end of an arch OFFC and either end of the loop, where the left and
right are determined by the direction parameter \( d^* \). It is opposite for the right loop. The method has been shown to be robust to a certain amount of rotation that can be expected in the realistic operational environment. It does not, as of now, allow one to determine the loop type correctly if the fingerprint image is submitted upside-down. The test for loop type should only be performed for those arches that show significant change of amplitude of \( \cos(\gamma) \) (close to 90 degrees). Values of \( \cos(\gamma) \) corresponding to low curvature OFFCs should not be admitted as they very likely correspond to insignificant OFFCs. Such pruning is possible, and it has been added for the purpose of these studies. The test for the OFFC type, including the removal of insignificant entries, can be performed using a zero-crossing counter with hysteresis.

A rule-based classifier algorithm based on the OFFC features can be constructed as follows:

```plaintext
if number_of_whorl_OFFCs>0
    category='whorl';
else if number_of_loop_OFFCs>0
    for each pair of loop - arch OFFC determine the number of left and right loop OFFC;
    if number_of_left_loop_OFFCs> number_of_right_loop_OFFCs
        category='left loop';
    else if number_of_left_loop_OFFCs< number_of_right_loop_OFFCs
        category='right loop';
    else
        category='loop';
    end
else if number_of_arch_OFFCs>0
    category='arch';
end
```

The two databases that present different levels of difficulty for the classification algorithms have been used to verify the implementation of the OFFC based classifier with the proposed modifications. The first database, FVC 2004 DB1b, is considered easier to classify than the FVC 2000 DB3b database. It is expected that processing techniques that are designed to smooth the orientation field would cause significant improvement of performance in the case of a low quality dataset. It should be noted that we are interested primarily in errors caused by the low quality of impressions. Unfortunately, some of the fingerprints are represented by partial images, which can lead to erroneous classification. Similarly, there are errors introduced as a result of the limitations of the classification method. Errors are particularly evident for some cases of tented arches (misclassified as left or right loops) and loops for which delta and core singularities are located.
close to one another (typically within 2 ridge-furrow periods). However, those challenging images have not been excluded from the tests presented here for the sake of objectivity. The results are presented in Table 4.5 below for four cases. The orientation field was extracted on a grid of 11 pixels. In cases 1 and 2 the block size was chosen as \( c_b = 19 \) pixels (with a 4 pixel overlap); the block size chosen in this case is based on numerous experiments and has been shown to provide good categorization performance. In case 1, no smoothing was used. In case 2, one iteration of orientation field smoothing was added using weighted averaging on a 3x3 neighborhood. The weights were obtained from a Gaussian function with standard deviation 3. Case 3 uses the block size selected according to the method presented in the previous section, where the block sizes are increased until a desired smoothness of the orientation field is achieved. Case 4 uses the algorithm proposed in the previous section, which is based on the MRF method on a grid of orientation estimates from 11x11 pixel blocks (with no overlap). In all cases, Gaussian noise \( N(0,0.0001) \) was added at the input in order to avoid division by zero in the algorithm used for estimation of variance (5-30). The categorization performance is reported in Table 2, and it can be concluded that for the lower quality dataset it is beneficial to apply the smoothing methods proposed herein. More extensive tests on different datasets would be desirable. In particular, it would be useful to define standardized datasets with assigned quality rankings.

<table>
<thead>
<tr>
<th>Recognition Rate 4/3 class [%]</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
<th>Case 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FVC 2004 DB1b</strong></td>
<td>87.50/88.75</td>
<td>86.25/88.75</td>
<td>87.50/90.00</td>
<td>87.50/88.75</td>
</tr>
<tr>
<td><strong>FVC 2000 DB3b</strong></td>
<td>55.00/57.50</td>
<td>68.75/68.75</td>
<td>62.50/62.50</td>
<td>72.50/75.00</td>
</tr>
</tbody>
</table>
CHAPTER 5

CERTAINTY OF ORIENTATION ESTIMATE

In this chapter it is shown how low quality influences orientation estimation. In Section 5.1, representative methods for orientation estimation are presented and the concept of orientation signature is introduced. In Section 5.2 the effect of low quality is discussed in the context of local orientation estimation. Observations and conclusions from Section 5.3 serve as the basis for the formulation of quantitative measures of quality based on uncertainty of local orientation estimates.

5.1 Orientation Estimation Methods

The important features for fingerprint categorization can be obtained from local ridge line orientations. In the previous chapter it has been shown how a silhouette of the fingerprint can be found in the process referred to as segmentation. This is useful in order to limit the required amount of processing. For instance, the feature extraction process is conducted only within the fingerprint silhouette. Important features such as SPs and OFFCs can be obtained from analysis of the orientation field, which consists of orientation estimates computed based on local information. The local orientations are typically computed on a grid \( \Lambda \) of image blocks \( B_\lambda \) and constitute the orientation field over the fingerprint image. The orientation field typically undergoes certain regularization operations and is subsequently used in extraction of higher level features, in image enhancement, and in classification. For details regarding the fingerprint categorization system building blocks refer to the previous chapter.

Four methods of orientation field estimation presented below include: oriented window method, Pseudo-Polar Fourier Transform (PPFT)-based method, Gabor filter-based, and PCA of gradients. They represent the methods commonly encountered in the area of automated fingerprint processing and classification. The oriented window is used for the estimation of local ridge frequency. The PPFT-based method is a variant of the Fourier Transform approach, and has not been used before for the purpose of the orientation estimation. The Gabor filter bank approach has been previously used for enhancement of fingerprint images.
In the course of this work it has been determined that it is better to remove the mean prior to applying any of the methods. Particularly useful for this purpose, and computationally efficient, is the gradient operator. Techniques used for fingerprint image normalization, which are encountered in the literature and involve removing the mean and image filtering, can be used to this effect [Maltoni et al., 2003].

In the oriented window method, the image is divided into a grid of square blocks $B_x$, possibly with an overlap. Subsequently, each block undergoes a rotation transformation and pixel values are summed up along the columns, which yields an orientation signature $s(\theta)$ (5-5). First, let

$$[x' \ y']^T = T(\theta) \cdot [x \ y]^T \tag{5-1}$$

where the coordinate system is rotated using rotation matrix $T$, with $\theta$ the rotation angle in the counterclockwise direction,

$$T(\theta) = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \tag{5-2}$$

$$I(x',y') \approx I(x \cos(\theta) + y \sin(\theta), -x \sin(\theta) + y \cos(\theta)) \tag{5-3}$$

$$(x,y) \in B_A \tag{5-4}$$

$I(x',y')$ denotes the rotated image, and the approximate equality in (5-3) is exact in the continuous case. Typically $(x',y')$ are not integers after the rotation transformation is applied to the integer coordinates $(x,y)$, and, as a result, the intensities $I(x',y')$ need to be obtained through interpolation. The orientation signature $s(\theta)$ is then defined as

$$s(\theta) = \text{Var} \left( \sum_{y'} I(x',y') \right) \tag{5-5}$$

The optimum orientation angle $\hat{\theta}$ is given by
\[
\hat{\theta} = \arg \max_{\theta} \{s(\theta)\} + \frac{\pi}{2}
\]

(5-6)

The underlying assumption behind this method is that the variance attains a maximum along the direction perpendicular to the local orientation of ridges. The main difficulty associated with this approach is that the rotation transformation requires interpolation of pixels, which alters data and, more importantly, it is computationally inefficient.

In the Fourier Transform-based approach, the image is also divided into a grid of blocks \( B \). In order to obtain an orientation estimate, it is convenient to work in polar coordinates. Then it can be shown that the orientation of maximum energy in the Fourier spectrum corresponds to the orientation perpendicular to the local ridge orientation. Instead of working with the polar Fourier representation, it suffices to use PPFT \([\text{Keller et al., 2005}]\). The properties of PPFT allow for estimation of energy at a given orientation in an easy way. The negative aspect of using PPFT stems from the fact that the frequency sampling along the radial direction is not equal at different angles and the usual filtering methods applicable to the DFT need to be modified. In comparison to Polar Fourier Transform approach, PPFT requires fewer computations, essentially equivalent to the Fast Fourier Transform (FFT), i.e. \( o(2\log_2 N) \) for an \( N \)-point sequence \([\text{Smith and Smith, 1995}]\). For a \( N \) by \( N \) pixel image, one needs to consider proper zero-padding as presented in \([\text{Keller et al., 2005}]\). The presence of noise affects the performance more significantly than does the Gabor filtering approach due to the presence of high and low frequency noise and spectral components that are due to anisotropic artifacts. Once the PPFT is calculated for a block, the orientation of maximum energy corresponds to the orientation perpendicular to fingerprint ridges in the block. The angular resolution is limited to

\[
\Delta \theta = \frac{\pi}{2c_b}
\]

(5-7)

where \( c_b \) is a block size.

For a typical value of \( c_b = 15 \) pixels (about 1-2 ridge-valley distance in 500 dpi images), the resolution is 6 degrees - a disadvantage with respect to the gradient method that allows working with a continuous interval of angles.
The implementation of the method requires interpolation that can be efficiently performed in the
frequency domain. It needs to be done twice for angular frequencies (5-8) and (5-9):

\[
f_{\theta_1} = \left( -\frac{\pi}{4}, \frac{\pi}{4} \right) \cup \left[ \frac{3\pi}{4}, \frac{5\pi}{4} \right]
\]

\[
f_{\theta_2} = f_{\theta_1} + \frac{\pi}{2}
\]

Only the first case is presented for brevity below. For a more detailed description see [Keller et
al., 2005], [Averbuch et al., 2006].

Step 1) Zero-append columns by adding N+1 rows

Step 2) Perform (2N+1) point DFT of each of N columns

Step 3) On each of 2N+1 rows of the resulting 2N+1 by N matrix perform a fractional Fourier
Transform:

\[
(F^a X)(l) = \sum_{u=-N/2}^{N/2} X(u) \cdot e^{-j2\pi a k l / N}, l = -N/2, ..., N/2
\]

Step 3 ensures constant angular frequency along each of the columns by properly selecting \( \alpha \) as a
function of a row number \( k \):

\[
\alpha = \frac{2k}{N}
\]

Assuming that the signal in the transform domain can be represented as \( R(\omega_\rho, \omega_\theta) \), where \( \omega_\rho \),
\( \omega_\theta \) represent angular frequencies in the radial and angular directions respectively. The orientation
signature can be defined as

\[
s(\theta) = \frac{1}{N} \sum_{\omega_\rho} |PPFT\{I(x, y)\}|^2, \omega_\rho > 0.
\]
From (5-12), similarly to the oriented window method, the local orientation $\hat{\theta}$ is found according to (5-6).

In the gradient method, the gradient field is first computed for the fingerprint image. For the sake of illustration, Sobel kernels are used. Next, the dominant orientation can be found by examining the distribution of gradient orientations denoted by $p(\angle g)$. This distribution can be approximated using Parzen windows for gradients in each block $B_\lambda$ while keeping in mind that it is a circular distribution. It has typically a multimodal form in which additional modes are due to the presence of oriented patterns other than ridges within $B_\lambda$. The maximum mode can be taken as an estimate of the orientation perpendicular to the local orientation of ridge lines. It corresponds to the dominant orientation in a maximum likelihood sense; however, it ignores the magnitude of the gradient vectors. It is preferred in practice to use PCA of the covariance matrix of gradients; its very efficient implementation has been recently proposed by [Bazen and Gerez, 2002]. Given the covariance matrix of the gradients,

$$\Sigma = \begin{bmatrix} G_{xx} & G_{xy} \\ G_{xy} & G_{yy} \end{bmatrix}$$  \hspace{1cm} (5-13)

where

$$G_{xx} = \frac{1}{|B_\lambda|} \sum_{\hat{g}} g_x^2$$ \hspace{1cm} (5-14)

$$G_{yy} = \frac{1}{|B_\lambda|} \sum_{\hat{g}} g_y^2$$ \hspace{1cm} (5-15)

$$G_{xy} = \frac{1}{|B_\lambda|} \sum_{\hat{g}} g_x g_y$$ \hspace{1cm} (5-16)

one can find principal vectors and principal values of $\Sigma$. The orientation of the second principal vector corresponds to the mean-squared error optimal ridge orientation estimate in block $B_\lambda$. Alternatively, one can use non-uniform weights under the summations in (5-14), (5-15) and (5-16) in order to obtain $\Sigma$ [Lindeberg, 1994], [Kass and Witkin, 1987], [Bazen and Gerez, 2002]. With respect to the other methods (oriented window, PPFT - based, and Gabor filter response - based, and maximum of the distribution of the gradient orientations), whenever gradient response in this approach is characterized by more than one mode of orientation, the
PCA of gradients does not find the dominant orientation (global maximum), but rather a mean orientation of gradients weighted by gradient magnitude.

![Figure 5.1](image)

Figure 5.1. Results of orientation field estimation using gradient method; a) block size is 17 (no overlap), b) block size is $c_b = 33$, grid is $c = 17$ (overlap of 8 points).

Figure 5.1 above shows the orientation field estimated on a grid $\Lambda$, where the distance between grid points $c=17$, using PCA of gradient vectors. In Figure 5.1 a), the orientation field was estimated on a grid of blocks without overlap for $c_b=17$. It can be seen that variations in the quality of the ridge pattern affect local orientation estimates and the resulting field is locally non-smooth in some regions. This undesired effect can be remedied by simply increasing the block size. The orientation field for the image b) has been estimated using the same grid $\Lambda$, but the block size is $c_b=33$. It can be noted that the orientation field is smoother.

The local orientation can also be estimated using matched filters; several such filters have been proposed in recent years [O’Gorman and Nickerson, 1989], [Sherlock et al., 1994]. Of particular interest is a set of biologically inspired Gabor filters. A bank of Gabor filters has been proposed for application to ridge enhancement by [Hong et al., 1998]. In the spatial domain the Gabor filter can be represented as a product of a Gaussian function, $f_{bp}$, and a sine wave, as shown below:

$$f_{bp}(x', y') = f_{bp}(x', y') \cdot e^{-0.5[(x'/\sigma_x)^2 + (y'/\sigma_y)^2]} \cdot \cos(2\pi \cdot f_0 \cdot x')$$  \hspace{1cm} (5-17)

where $(x', y')$ are obtained after applying the rotation transformation (5-1), and $\sigma_x$, $\sigma_y$ are shape parameters. The Gabor filter has bandpass properties, and the band is centered around $\pm f_0$. The shape is determined largely by the shape of the Gaussian function (in practical applications it is...
true up to the effect of leakage between the DFT harmonics due to the finite size of the analysis window). This effect can be limited by appropriate selection of the window size with respect to the shape parameters. The bandpass response function $F_{bp}(u,v)$ is given by

$$F_{bp}(u,v) = .5 F_{lp}(u,v) \ast \left( \delta(u + u_o, v + v_0) + \delta(u - u_o, v - v_0) \right),$$  \hspace{1cm} (5-18)$$

where $F_{lp}(u,v)$ is the Fourier Transform of the Gaussian function $f_p(x',y')$. If $F_{bp}$ is the Fourier Transform of the Gabor filter and $FT\{I(B_2)\}$ is the Fourier Transform of the pattern $I(B_2)$, then the orientation signature can be defined as

$$s(\theta) = \sum_{u,v} |F_{bp}(u,v) \cdot FT\{I(B_2)\}|^2$$  \hspace{1cm} (5-19)$$

and the dominant orientation is obtained using (5-6), and signifies the orientation of maximum energy in the spatial-frequency band centered around $(u_o, v_o)$ and $(-u_o, -v_o)$. Note the similarities between this method and PPFT and oriented window methods.

### 5.2 Orientation Estimation in Presence of Defects

In order to better understand the impact of image quality on fingerprint processing and classification, we would like to focus on local orientation of ridges. Ridge orientation as a feature is commonly used and has application in fingerprint categorization, image enhancement, minutiae and singularity detection. Similar techniques as shown herein are widely used beyond the area of automated fingerprint processing as well, for instance in analysis of oriented patterns [Kass and Witkin, 1987]. In fingerprint classification, an important role is associated with a ridge flow orientation field. In the process of deriving an orientation field representation, one often proceeds by dividing the image into a grid of blocks. Within each block, orientation estimation is performed and it is then combined with other block estimates to form an orientation field. Certain regularization techniques are often used in order to improve the estimate. Frequently such techniques are based on the assumption of local smoothness that, as it can be seen from sample orientation fields, is not valid in general, but has been shown to produce very encouraging results. After presenting results of experiments and their discussion, a definition of quality is proposed.

In Chapter 3, sample images from fingerprints sensors have been shown. They have been selected to provide coverage of various types and various amounts of quality defects. In this section the
results of processing selected regions of those images are presented. The goal of this section is to provide an experimental basis on which theoretical analysis can be formulated. These specific cases help facilitate understanding of the impact that certain patterns with and/or without quality defects have on orientation estimation. After inspection of a large number of images, the following defect modes have been identified as the most important: 1) weak anisotropy, 2) presence of more than one anisotropic pattern, one of which corresponds to a spurious artifact. The case 1) may, for instance, originate in the presence of a high-variance random noise, and is often encountered in a low contrast area of a fingerprint image. The case 2) is most commonly associated with skin folds or scars. The detailed model of defects is presented in Chapter 3. For now, let us associate a good quality pattern with the one that is anisotropic, i.e. it has one dominant orientation. This justifies using scalar measures of the uncertainty of the orientation estimate. The cases involving the orientation signature are presented in this section.

The selected examples from natural impressions presented in Figure 5.2, Figure 5.3, Figure 5.4 and Figure 5.5 illustrate different degrees of the defects. The figures (clockwise from top left) present orientation signatures computed from a) PPFT, b) oriented window, c) Gabor filter bank methods, and d) scatter plot of gradients, e) histogram of gradient orientations. In all cases, except for the gradient scatter-diagram, the orientation signatures have been normalized. The magnified block of image is shown as well. The PCA of gradient vectors is considered to be the fastest method. However, other techniques such as Fourier Transform-based, the variant of which is PPFT-based and was presented in the previous section, matched filter-based (Gabor filter bank) and spatial (projections onto rotating axis) are also encountered. Some of the properties of these methods are best learned from visual inspection, and that is one of the goals of this section. They are followed by quantitative analysis further in this and the following chapter.

The case of a good quality block has been presented in the Figure 5.2. All methods perform rather well. The PPFT signature contains a wide spike around zero degrees, which is explained by the non-constant orientation. The curvature increases toward the lower part of the image. It is similar in the case of the oriented window method. The Gabor filter bank response is characterized by a wide smooth spike — the result of the bandpass character of the matched filter. The spike width is related to filter's bandwidth and the curvature of the fingerprint pattern. The smoothness can be attributed to the overlap between the rotated filters in the bank. The asymmetry can be observed in the PPFT histogram of gradient orientations and Gabor filter bank cases, and is very weakly manifested in the case of the rotated window method. The asymmetry is not easily explained by visual inspection but it may indicate an oriented component at the angle between 10 and 20
degrees. It can be seen that the histogram of gradient orientations contains spikes not consistent with the dominant orientation and is, as it will be shown later, too sensitive to quality defects to be directly used in orientation estimation. The scatter-plot of gradients exhibits a characteristic ‘butterfly’ shape due to the curved shape of ridges indicating proximity of a SP. This effect induces higher uncertainty of the orientation estimate since all of the orientation estimation methods assume locally constant orientation of ridges.

Figure 5.2. Orientation signatures for a good quality pattern (impression 103_1 from FVC 2004 DB3b).

Figure 5.3 depicts the case of a low contrast sample with noisy artifacts. The quality is very low and it is rather difficult to distinguish the ridge-valley pattern by inspection. However, there is a pattern present that is oriented at slightly more than 45 degrees measured counter-clockwise with respect to the horizontal axis. Methods based on oriented window, PPFT and Gabor filter bank
were able to detect an oriented pattern. Naturally, the low quality of the pattern affects the orientation signatures by decreasing the signal-to-noise margin. This is evident in the histogram of gradient orientations and the scatter diagram of gradients, where the noise dominates the oriented pattern.

Figure 5.3. Orientation signatures for a low quality pattern (impression 105_6 from FVC 2000 DB3b).

The pattern shown in Figure 5.4 contains over-emphasized ridges. Nevertheless, in spite of the visual appearance, the orientation estimation can be accomplished very reliably – see especially PPFT, window and Gabor filter-bank responses. The three spikes visible in the histogram of gradient orientations may have been introduced by the sensor (they could be observed for multiple images from the same sensor).
Figure 5.4. Orientation signatures for a sample containing over-emphasized ridge pattern (impression 102_8 from FVC 2004 DB3b).

Figure 5.5 shows the sample image and the orientation signatures for the image with a defect resembling a crease. Such a defect introduces a strong peak, which is clearly seen in the case of PPFT, window and histogram methods. In the case of PPFT-based and oriented window methods, the defect produced a stronger response than the actual ridge pattern, thus leading to erroneous orientation estimation. The Gabor filter bank, on the other hand, shows response to the ‘true’ orientation of the periodic pattern of ridge-valleys and ignores the defect, which is clearly advantageous.
Figure 5.5. Orientation signatures for a ridge pattern containing directional defect (impression 108_1 from FVC 2000 DB3b).

It can be concluded that the most severe effect on the orientation estimates stem from 1) a high level of noise and 2) an oriented artifact, such as a crease, that has size comparable with the ridge-valley period and is not tangent to local orientation. Table 5.1 below shows the local orientation estimates for different images shown in the figures above and for different methods.
Table 5.1. Estimated local ridge orientations for the patterns shown in Figure 5.2 - Figure 5.5 and for 5 different estimation methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>Figure 5.2</th>
<th>Figure 5.3</th>
<th>Figure 5.4</th>
<th>Figure 5.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPFT</td>
<td>0</td>
<td>51.77</td>
<td>-67.01</td>
<td>6.91</td>
</tr>
<tr>
<td>Window</td>
<td>0</td>
<td>49.09</td>
<td>-65.45</td>
<td>10.91</td>
</tr>
<tr>
<td>Matched filter</td>
<td>2.73</td>
<td>54.55</td>
<td>-65.45</td>
<td>49.09</td>
</tr>
<tr>
<td>PCA</td>
<td>2.56</td>
<td>36.71</td>
<td>-66.22</td>
<td>43.87</td>
</tr>
<tr>
<td>Mean orientation</td>
<td>3.24</td>
<td>38.40</td>
<td>-54.92</td>
<td>37.58</td>
</tr>
</tbody>
</table>

The interpretation of those results requires additional information such as the variance or uncertainty of the estimate. Based on numerous experiments, it has been concluded that the Gabor filter-bank approach presents a robust way of estimating oriented patterns and is shown to have superior performance. The oriented window shows similar performance to the PPFT-based method. In the examples presented above, the PPFT-based method is in fact very close in performance to the Gabor filter-bank method in that it is able to detect the multiple orientations. The oriented window method seems to have lower resolution with respect to multiple oriented patterns. Perhaps this result can be explained by the low pass character of the interpolation procedure. The advantage of the Gabor filter-based method reveals itself in the case of presence of an anisotropic defect. The crease pattern introduces high and low frequency components to the spectrum, which are filtered out in the Gabor matched filtering approach so that only one strong maximum of the orientation signature is present. We postpone the discussion of PCA of the gradient field until later. The direct comparison presents certain problems that can be alleviated to a certain extent by introduction of the orientation uncertainty.

We conclude by defining first the quality of the local orientation estimate and secondly show how local quality information can be used to quantify the quality of the fingerprint image. As it can be seen from the examples presented above in Figure 5.2 through Figure 5.5, the ridge orientation is estimated within a local window. This is done by an operation that results in an orientation signature. Based on the signature, one can select the most plausible orientation, which corresponds to the orientation for which the orientation signature has its maximum. Clearly this is not the only possible choice of orientation, and there is a finite probability that the true local ridge orientation corresponds to one of the local maxima.
Suppose that we have an image pattern \( I \) defined over the grid of \( \Lambda \), \( c=1 \) and a set of features \( x_\lambda \in X \) extracted at locations \( \lambda \in \Lambda \), such that \( x_\lambda \) is defined for each block \( B_\lambda \) at location \( \lambda \).

Let \( G \) be a function (feature extraction) defined on a block \( B_\lambda \) such that:

\[
x_\lambda = G(I(B_\lambda))
\]

This is a hard assignment – for a given pattern \( B_\lambda \), there is one orientation that, according to a given orientation estimation method, is the best choice. In order to compute the quality of a pattern \( B_\lambda \), one can employ a soft feature extraction instead, as discussed in Chapter 2, Eqn. (2-8). In this approach

\[
s(x_\lambda) = F(I(B_\lambda))
\]

which reflects the fact that there is a finite probability that the feature \( x_\lambda \) is anywhere in the feature space \( X \). The quality can be then defined as a function \( f \) of the feature signature \( s(x_\lambda) \), that is, ideally, a probability density function \( p(x_\lambda) \) that expresses the uncertainty of the feature \( x_\lambda \) that reflects quality of \( B_\lambda \). In the case of fingerprint ridge orientation estimation, \( F \) can be one of the methods presented in Section 5.1 and \( s \) in this context is an orientation signature, where the feature \( x_\lambda = \theta_\lambda \), \( \theta_\lambda \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \).

The quality of the pattern in \( B_\lambda \) given the feature signature is

\[
q_\lambda = f \circ s, \quad q_\lambda \in [0,1] \tag{5-20}
\]

Now, suppose that we have an image \( I_A \) and that we are interesting in finding its quality. Suppose \( q_{A,\lambda} \), the quality measure at \( B_{A,\lambda} \) (block \( B_\lambda \) of image \( A \)), is a random variable, then the probability of the following event:

\[
q_{A,\lambda} > q_{th}
\]

where \( q_{th}=0.5 \) is the quality threshold, can be used as a global quality metric for the image \( I_A \) as shown in equation (5-21).
\[ P(q_{A,\lambda} > q_{th}) = q_A \] (5-21)

Equation (5-21) can be interpreted as the proportion of the blocks \( B_{A,\lambda} \) where the quality is greater than \( q_{th} \). Any two patterns \( I_A \) and \( I_B \) with the sets of quality measures \( q_{A,\lambda} \in Q_A \) and \( q_{B,\lambda} \in Q_B \), respectively, can be then compared using the following approach:

\[ P(q_{A,\lambda} > q_{B,\lambda}) = q_{AB} \]

\( q_{AB} \) can serve as a comparative quality metric, such that

- \( q_{AB} > 0.5 - I_A \) is of better quality than \( I_B \)
- \( q_{AB} = 0.5 \) - \( I_A \) and \( I_B \) have similar quality
- \( q_{AB} < 0.5 \) - \( I_B \) is of better quality than \( I_A \)

It is useful if one has two impressions of the same finger and it is desired to select the one better suited for processing. Additional precautions are needed. The quality alone, as defined herein, does not carry information about the classification outcome. For instance, a good quality pattern of ridges with insufficient number of minutiae will likely lead to a classification error. Nevertheless, there is a benefit from using the localized quality descriptors. Using a solely scalar metric can lead to loss of information. The spatially localized vector metric allows for more flexibility. For instance, one can associate a quality metric with different salient features, such as minutiae points of a fingerprint. The quality of two impressions can then be compared on the basis of the metric associated with minutiae. One can conceive numerous other examples in which different features and different classification problems can be considered. [Tabassi et al., 2004] used a vector of quality features including a count of minutiae points found in different quality regions. It was subsequently used to train a neural network classifier and used to predict the classifier’s performance. It appears that the quality metric should be designed in connection with the processing techniques used. This should lead to a better ability to predict the classification outcome. For this reason, the methods of deriving uncertainty for different orientation estimation methods are proposed in the next section.
5.3 **Orientation Uncertainty – Quantitative Approach**

It has been previously mentioned that a quality metric for images used in pattern recognition should reflect the uncertainty associated with extracted features. This approach is attractive, because once the uncertainty is known, it can be used in higher level feature extraction. For instance, the information from high quality regions can be propagated and used to reinforce our knowledge in the neighborhoods of low quality. A low quality image, according to our definition, is associated with high uncertainty of features. One way of propagating information is through orientation diffusions. As numerous examples show [Perona, 1998], such an approach gives surprisingly good results. In the case of fingerprints that have highly oriented and periodic character, the propagation of the information can be directed by both local quality and by the orientation of the patterns in local neighborhood.

In the deterministic approach, one seeks ways to evaluate the quality of orientation as a deterministic function of a signal, which in our case is a two-dimensional pattern. This approach is commonly used, for example, where the function can have the form of a coherence measure. It can be defined directly as a function of gradients or, equivalently, as a function of eigenvalues of the covariance matrix of gradients. The definition of coherence ($c_1$ and $c_2$) can be found in Appendix A. It has a statistical interpretation – it is related to a measure of dispersion for circular distributions. It can assume values from 0 to 1. It is zero for patterns in which intensity gradients are distributed uniformly in angle and/or cancel out. It is unity for the cases in which gradient vectors are parallel. Some of the properties of coherence have been pointed out by [Jiang, 2005]. In particular, the presence of relatively few strong parallel gradients can result in coherence close to 1 even if the majority of gradients have low magnitude and are due to low quality. It can be avoided when coherence is computed on normalized gradients. As a result of normalization, all gradients have equal weights. As shown below, most of the time the normalization results in decrease of coherence (see also Appendix A). We present an image block from the boundary of the fingerprint. It contains a noisy region of background and the fragment containing ridge-valley structure.
Figure 5.6. Boundary of background and fingerprint ridge pattern; coherence of gradients is high.

The gradient response is plotted in the form of scatter diagrams and histograms of gradient orientations, as shown below in Figure 5.7.

Figure 5.7. Figure shows scatter diagrams of gradients: for the entire image (Figure 5.6) a), bottom right quadrant with the ridge pattern b), and the top left quadrant containing background noise c). The diagrams d), e) and f) depict corresponding gradient orientation histograms.

Table 5.2. Comparison of the coherence of normalized $c_2$ and non-normalized gradients $c_1$ for the cases presented in Figure 5.7.

<table>
<thead>
<tr>
<th>Region</th>
<th>$c_1$</th>
<th>$c_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ridge region</td>
<td>0.79</td>
<td>0.45</td>
</tr>
<tr>
<td>Background region</td>
<td>0.34</td>
<td>0.10</td>
</tr>
<tr>
<td>Ridge and background region</td>
<td>0.79</td>
<td>0.36</td>
</tr>
</tbody>
</table>
From the table it can be seen that the coherence is lower after normalization. It confirms that the strong gradients cause the coherence of squared gradients $c_1$ to be high despite of the presence of the large number of weak gradients corresponding to the background region. The coherence $c_1$ evaluated in a good quality area is high – in this case, it is equal to the coherence $c_1$ computed for the entire block. Unfortunately, normalization caused significant decrease in coherence $c_2$ in the case of the entire block and in the case when the block contains the ridge-valley region. It leads to the conclusion that $c_2$ may be oversensitive to the noise. A certain amount of noise present in the ridge-valley area gives rise to weak gradients. The normalization of gradients increases the influence of weak gradients on $c_2$. On the other hand, $c_2$ has lower value in the case of the entire block than in the case of the ridge-valley region. Such sensitivity is a desired feature of the quality metric and it may be interesting to investigate further.

In spectral, matched filtering and oriented window methods the approach to orientation estimation proceeds as follows. For each of the cases we define the orientation signature $s(\theta) = f(I(B_\theta))$, $\theta \in \Theta$, which is a result of an operation $\phi$ applied in the analysis window $B_\theta$. The best candidate orientation is then selected according to (5-6), where $\hat{\theta}$ is the angle at which $s(\theta)$ has the strongest response, or equivalently the orientation along which the variation of the pattern is the greatest. Typically, $s(\theta)$ has a number of local maxima and the orientation that is selected corresponds to the strongest one. A quality measure $q_\theta$ for each block $B_\theta$ can then be constructed according to (5-20), in which the following conditions need to hold for $f$:

1. Invariance to linear scaling of $s(\theta)$:
   
   \[ f(s(\theta)) = f(ks(\theta)), \quad k \in \mathbb{R} \]

2. Increasing function of the signal-to-noise margin:
   
   \[ f(s(\theta)) \sim |s(\hat{\theta}) - s(\theta)| \]

3. Decreasing function of the signature value $s(\theta)$ weighted by the distance from the orientation estimate:
   
   \[ f(s(\theta)) \sim \frac{1}{s(\theta) \cdot (\hat{\theta} - \theta)^2} \]

4. The values of $f(s(\theta))$ should be contained in the closed interval $[0,1]$

One equation for $f(s(\theta))$ that fulfills all of the above requirements is the following form:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
where \( \hat{\theta}, \theta \) are in radians. Let \( M \) be the set of all points at which \( s(\theta) \) is a local maximum. Further let \( M_{\hat{\theta}} \) be the set of all points at which local maxima are equal to \( s(\hat{\theta}) \)

\[ M_{\hat{\theta}} = \{ \theta : s(\theta) = s(\hat{\theta}), \ \theta \in M \} \]

The maxima need to be distinct so that points that belong to one weak maximum are counted as 1. In particular, if \( s(\theta) \) is constant, then \(|M| = 1\). The following definitions clarify the concept of strong and weak maximum.

For a point \( s(\theta) \) to be a strong local maximum \( M \), we need to verify that there exists \( \varepsilon \), such that for any \( \alpha < \varepsilon \) the set

\[ \{ \theta : s(\theta) > s(\theta + \varepsilon), \ s(\theta) > s(\theta - \alpha) \} \]

is non-empty,

and for the weak maximum the set

\[ \{ \theta : s(\theta) \geq s(\theta + \varepsilon), \ s(\theta) \geq s(\theta - \alpha) \} \]

is non-empty,

where we need to verify that \( s(\theta + \varepsilon) \) and \( s(\theta - \varepsilon) \) are also weak maxima.

It can be easily verified that \( f(s(\theta)) \) in the equation (5-22) equals at most one when at most one local maximum is present in \( s(\theta) \). Consequently, in the cases when \( \hat{\theta} \) cannot be determined because \( |M_{\hat{\theta}}| > 1 \), then the maximum value achievable by \( f(s(\theta)) \) is \( 1/|M_{\hat{\theta}}| \). The non-zero response at any \( \theta \not\in M_{\hat{\theta}} \) further decreases \( f(s) \). The amount that each point \( s(\theta) \) contributes to the denominator of \( f(s(\theta)) \) is weighted by the normalized squared distance from \( \hat{\theta} \) to \( \theta \) and by

\[
f(s) = \frac{1}{|M_{\hat{\theta}}| + \frac{4}{\pi^2} \sum_{\theta \in \Theta \setminus M_{\hat{\theta}}} \frac{s(\theta) \cdot (\hat{\theta} - \theta)^2}{|s(\hat{\theta}) - s(\theta)|}}
\]
\[
\frac{1}{|s(\theta) - s(\theta)|},
\]
which implies low values for the cases of peaked and unimodal \(s(\theta)\). As can be seen from (5-22), the scaling of \(s(\theta)\) does not influence the value of \(f(s)\).

However, a small modification is necessary. Note that for the cases of wide maxima, the value of \(f(s(0))\) underestimates the quality as shown in the figure below.

![Figure 5.8. Orientation signature obtained from the response of Gabor filters a) and the corresponding quality descriptors \(f(s)=0.029\) and \(f'(s)=0.353\). The corresponding fragment of a fingerprint image b).](image)

Figure 5.8. Orientation signature obtained from the response of Gabor filters a) and the corresponding quality descriptors \(f(s)=0.029\) and \(f'(s)=0.353\). The corresponding fragment of a fingerprint image b).

The improvement with respect to the proposed measure of \(f(s(\theta))\) can be achieved by including under the sum in the denominator of (5-22) only the points representing local maxima of \(s(\theta)\):

\[
f'(s) = \frac{1}{|M_2\| + \frac{4}{\pi^2} \sum_{\theta \in M_2 / M_2} \frac{s(\theta) \cdot (\hat{\theta} - \theta)^2}{s(\theta) - s(\theta)}}
\]

(5-23)

The value of \(f'(s(\theta))\) computed using the equation (5-23) is much higher as can be seen in the Figure 5.8. It is due to the fact that it acts only in a discrete number of locations that correspond to maxima of \(s(\theta)\) and, as a consequence for non-smooth curves with frequent local maxima, \(f'(s(\theta))\) has low values.

It has been found that better results yet can be obtained when the following definition is used:
Equation (5-24) is motivated by the fact that the most of the mass of $s(\theta)$ should be concentrated around the estimate $\hat{\theta}$. Consequently, the ratio in (5-24) is close to 1 when the orientation signature is peaked around $\hat{\theta}$. It quickly approaches zero when $s(\theta)$ tends to more uniform shape. It can also be noted that when $s(\theta)$ is uniform, $f(s(\theta))=0$. The function (5-23) is always non-negative, which is not the case for (5-24). However, (5-24) is not as sensitive to the lack of smoothness of $s(\theta)$ as (5-23) is. This is important, for example, in the case of the orientation signature obtained from the Fourier spectrum.

Given the probabilistic models of the underlying data $I(B_\theta)$, one can quantify the uncertainty of the orientation signature using probability. It is reasonable to assume that there is a finite probability that the true orientation corresponds to $\theta \neq \hat{\theta}$. This uncertainty can be represented as a probability function. The approach below is valid for the case when $\theta$ is discrete. The probability distribution of the orientation estimate $\hat{\theta}$ can be written as follows:

$$P(\hat{\theta}) = P(s(\theta) = \max \{ s(\theta) \} )$$  \hspace{1cm} (5-25)

In a discrete case we have $s_i = s(\theta_i), i=1,2,...,N. s_i$ can be viewed as a random variable for which we have the cumulative distribution function (cdf) $F_i(s_i) = P(s_i \leq s_i)$. We will also make an assumption of independence of random variables $s_i$. Then the probability that $s_i$ is the actual maximum, and $\hat{\theta} = \theta_i$, can be written as

$$P(s_i = \max) = \int \int ... \int p(s_1,...,s_N)ds_1...ds_N = \int \left( \prod_{i \neq l} \int p(s_i)ds_i \right) ds_i =$$

$$= \int p(s_j) \left( \prod_{i \neq l} F_i(s_j) \right) ds_j$$  \hspace{1cm} (5-26)
It expresses the probability that a given $s(\theta)$ is a global maximum. In practice, the assumption of independence may not always hold. For instance, the orientation signature obtained based on the responses of the Gabor filters is one such case. However, when equation (5-26) otherwise holds, it can be implemented for small $N$. For large $N$ it is better to use log-likelihood, but even then it becomes rather tedious to evaluate. The results obtained using this method are illustrated in the Figure 5.9 below.

Figure 5.9. Signal $x(n)$ a), and variance of frequency localization b).

The signal used in the example in Figure 5.9 a) has the following form:

$$x(n) = \sin(2\pi \cdot 0.04 \cdot n)(1 - h(n - 50) + h(n - 67)) + 0.5 \sin(2\pi \cdot 0.24 \cdot n)$$

where $h(n)$ denotes a discrete step function. Given the Fourier spectrum of $x(n)$, the task is to find the carrier frequency $f_0=0.04$. The signal $x(n)$ (see Figure 5.9 a) is corrupted by Gaussian noise, an interfering signal with normalized frequency $f_1=0.24$, and is absent for a duration of 17 samples. The task of finding the frequency can be accomplished by finding the maximum in the power spectrum. The figure depicts the signal with the noise and the plot of the sample and predicted variance estimates for a changing level of the input noise (expressed as a variance of i.i.d. zero-mean Gaussian noise on the horizontal axis). It can be noted that the predicted variance roughly follows the sample estimate.

Given $P(\hat{\theta})$, one can also use the entropy as a measure of uncertainty of the orientation estimate. In this case the good quality corresponds to low values of the entropy $f^{IV}(s(\theta))$, where
In the case of a continuous variable \( \theta \) with probability density function \( p(\theta) \), differential entropy can be used.

For the cases in which the probability function \( P(\hat{\theta}) \) cannot be found, however, the \( s(\theta) \) is available, given that \( \sum_{\theta \in \Theta} s(\theta) = 1 \), which can be ensured by normalization. When \( s(\theta) > 0 \), then (5-27) can be used to describe the quality, but it then will not have the usual information-theoretic interpretation.

In the case of orientation estimation using gradients of the intensities, the problem can be posed as a line fitting. The approximate formula for error propagation that has been presented by [Haralick, 1994] can then be used.

Let \( F(g, \theta) \) be a MSE function to be minimized, \( g = (g_x, g_y) \) is the gradient calculated for \( I(B_\theta) \), and \( \theta \) is the orientation angle between the x-axis and the line \( Ax + By + C = 0 \). Given that the statistical properties of \( g \) are known, in particular that the covariance matrix of \( g \) is known, we have that

\[
\sigma_{\Delta \theta}^2 = \left( \frac{\partial^2 F}{\partial \theta^2} \right)^{-1} \sum_{\Delta g} \left( \frac{\partial^2 F}{\partial g^2} \right)^T \left( \frac{\partial^2 F}{\partial \theta^2} \right)^{-1} \right)^T
\]

The detailed derivation of (5-28) is presented in Chapter 6. Sample results obtained for a set of synthetic images are presented in Figure 5.10.
Figure 5.10. Variance of orientation estimates; (x) – sample variance, (□) – predicted using (5), and (◇) function of coherence of gradients.

Figure 5.10 shows an experimental plot of sample and predicted variances, and a function of coherence, versus noise variance. Each data point was obtained from an 8-bit gray scale image, 128x128 pixels, oriented at -8 degrees, with zero mean Gaussian noise and variance from 0.001 to 10. Variance estimates were calculated on a grid of blocks of equal size (21x21 pixels) without overlap. It has been determined that the accuracy of the predicted orientation estimate variance, measured with respect to actual sample variance, decreases with increasing Gaussian noise.

The noisy synthetic ridge pattern (a sine wave oriented at 45 degrees with amplitude 32 and period of 11 pixels) is shown below in Figure 5.11.

![Test image distorted by additive Gaussian noise (with zero mean) with variance equal to 0.001 a), 0.01 b) and 0.1 c).](image)

Summarizing, we have proposed new methods for quantifying the uncertainty of the orientation estimate. Ideally, one should seek a probabilistic description. In realistic scenarios the probability
models are often not known. For the method using PCA of gradients of the intensity, the natural measure of uncertainty is known as coherence. It can be interpreted as a measure of how consistent the orientation of gradients in the neighborhood is. In the case of other methods, where the orientation corresponds to the maximum of the orientation signature, we have proposed alternative measures (5-23) and (5-24). The probabilistic approach in the form proposed in this section (5-25) and (5-26) is practical if the noise model of the data is known and if it can be written in a closed form. This is the case in the example shown in Figure 5.9, where under the assumption of Gaussian noise the points $s_i$ become $\chi^2$ random variables (under an approximate condition of independence). The remaining approach to estimating variance of the MSE estimate of orientation involving the error propagation formula (5-28) is also rather complicated. Instead, it may be more convenient to use the function of coherence to approximate the variance of the estimate of orientation.
CHAPTER 6

ANALYSIS OF THE EFFECT OF QUALITY DEFECTS ON ORIENTATION ESTIMATE

In Section 6.1 it is shown that the four representative methods of orientation estimation can be carried out in the Fourier domain. Following this notion in Section 6.2 and Section 6.3, the influence of additive noise and anisotropic artifacts is discussed and the closed-form solutions are presented for the orientation signatures obtained from the Fourier spectrum. The analysis is carried out for the continuous spatial case. Subsequently, the effect of a finite, discrete grid analysis window is discussed in Section 6.4.

6.1 Principle of Ridge Orientation Estimation

Local ridge line orientations are important features used in fingerprint recognition. The methods for their estimation have been presented in the previous chapter. It has also been shown how one can calculate the orientation uncertainty based on the orientation signature. For a given method and a local pattern that contains an oriented quasi-periodic pattern and noisy artifacts, one can measure how reliable the estimate is. In this chapter the performance of different methods in the presence of defects is examined. The analysis can be accomplished based on the models presented in Chapter 3. In order to simplify analysis, it is useful to point out that all of the methods are based on the same underlying principle – energy of the 2-D signal. Consequently, it suffices to demonstrate how defects influence the power spectrum. We have the following 4 methods: PPFT, spatial window, PCA of gradient, and Gabor filtering approach. Within the four methods, PPFT, spatial window and Gabor filtering (and any filter-based approach) seek the orientation of maximum energy. PCA of gradients gives a MSE-optimal estimate of the orientation.

Below we show that all of the methods can be related to the Fourier domain. Given the Fourier domain representation, a uniform approach can be taken. We need to show that the spatial window method and PCA of gradients can be equivalently carried out in the Fourier domain1.

---

1 The equivalency holds for the continuous case.
A gradient representation of the pattern allows for efficient orientation estimation. Let

\[ g(x, y) = \begin{bmatrix} g_x(x, y) & g_y(x, y) \end{bmatrix} \] \hspace{1cm} (6-1) \]

denote the gradient of the signal \( s'(x, y) \) that contains an oriented pattern. The selected orientation is normal to the orientation for which the mean squared value of the integral of residuals \( \delta \) (refer to the equation (6-2) below and Figure 6.1) is minimized.

\[ F(A, B, g) = \int \int \delta^2 dx dy = \int \int (A \cdot g_x(x, y) + B \cdot g_y(x, y) + C)^2 dx dy \hspace{1cm} (6-2) \]

\[ A = -\sin(\theta), \hspace{0.5cm} B = \cos(\theta), \hspace{0.5cm} C = 0 \]

\[ F(\theta, X) = \int \int (-\sin(\theta) \cdot g_x(x, y) + \cos(\theta) \cdot g_y(x, y))^2 dx dy \]

A, B and C are the parameters of the line equation. For simplicity, C is assumed to be 0. This parameter determines the displacement of the line from the origin and is irrelevant in the orientation estimation. Figure 6.1 illustrates the scatter diagram of gradient \( g \) (discrete case). The line \( Ax + By + C = 0 \) is oriented so that \( F \) is minimized. The arrows illustrate the residual errors.

\[ \frac{\partial F}{\partial \theta} = \sin(2\theta) \cdot (G_{xx} - G_{yy}) - 2 \cdot \cos(2\theta) \cdot G_{xy} = 0 \hspace{1cm} (6-3) \]

Figure 6.1. MSE fit of the line \( Ax + By + C = 0 \) to the data.

The optimum orientation \( \hat{\theta} \) corresponds to the root of the first derivative of \( F \) with respect to \( \theta \)
where \(G_{xx}, G_{yy}, G_{xy}\) correspond to quantities defined for the discrete case in the previous chapter; for the continuous case

\[
G_{xx} = \int \int (g_x(x,y))^2 \, dx \, dy \quad (a)
\]

\[
G_{yy} = \int \int (g_y(x,y))^2 \, dx \, dy \quad (b)
\]

\[
G_{xy} = \int \int g_x(x,y) g_y(x,y) \, dx \, dy \quad (c)
\]

In the equations (6-4) we ignore the integration interval and the windowing effect for the sake of simplicity. Assume the integrals are from \(-\infty\) to \(\infty\). The orientation that minimizes \(F\) is given by

\[
\hat{\theta}_q = 0.5 \cdot \text{atan} \left( \frac{2G_{xy}}{G_{xx} - G_{yy}} \right) \quad (6-5)
\]

It is the same as the result of the PCA conducted over the gradients. Since the PCA of the gradient vector field finds an orientation of gradients that is optimal in a MSE sense, it is sensitive to outliers. In particular, in the presence of more than one oriented pattern, the estimated orientation \(\theta\) does not correspond to either true orientation. Instead, it reflects a mean estimate weighted by the relative amplitudes of the gradients.

It can be shown that the orientation estimation using PCA leads to the same orientation estimate that results from analysis in the Fourier domain. Consider the gradient \(g(x,y)\) of the signal \(s'(x,y)\)

\[
g(x,y) = \begin{bmatrix} g_x(x,y) & g_y(x,y) \end{bmatrix}^T = \begin{bmatrix} \frac{\partial s'(x,y)}{\partial x} & \frac{\partial s'(x,y)}{\partial y} \end{bmatrix}^T. \quad (6-6)
\]

Assume that \(g(x,y)\) has a Fourier Transform equal to \(G(u,v)\). Then from the properties of the Fourier Transform we have

\[
G(u,v) = \begin{bmatrix} G_x(u,v) & G_y(u,v) \end{bmatrix}^T = \int [uS'(u,v) \quad vS'(u,v)]^T \quad (6-7)
\]

Now equations (6-4) written in the spatial domain can be equivalently represented in the Fourier domain.
\[ G_{xy} = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} u S'(u,v) v \left( S'(u,v) \right) ^* dudv = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} u |S'(u,v)|^2 dudv \]  

(6-8)

\[ G_{xx} = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} u S'(u,v) u \left( S'(u,v) \right) ^* dudv = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} u^2 |S'(u,v)|^2 dudv \]  

(6-9)

\[ G_{yy} = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} v S'(u,v) v \left( S'(u,v) \right) ^* dudv = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} v^2 |S'(u,v)|^2 dudv \]  

(6-10)

which follows from the Parseval's theorem. After substituting (6-8), (6-9), and (6-10) into (6-5) we have

\[ \hat{\theta} = 0.5 \cdot \text{atan} \left( \frac{2 \int_{-\infty}^{+\infty} u |S'(u,v)|^2 dudv}{\int_{-\infty}^{+\infty} u^2 |S'(u,v)|^2 dudv - \int_{-\infty}^{+\infty} v^2 |S'(u,v)|^2 dudv} \right) \]  

(6-11)

Interestingly, the same result as (6-11) is obtained from the power spectrum of \( s'(x,y) \). It follows from the observation that the integrals in (6-11) express second moments of the power spectrum of \( s'(x,y) \). We can therefore conclude that the optimum orientation estimate obtained from the MMSE criterion applied to the gradients and the MMSE applied to the power spectrum of \( s'(x,y) \) are in fact equivalent – they both find the orientation of maximum energy in the case of a pattern with a single orientation, or the MMSE optimum orientation in general (cf. PPFT-based method of orientation estimation, where the case of a pattern with a single orientation corresponds to the pattern for which the orientation signature has a single maximum).

The sensitivity of this estimate to small variance noise can be examined using the error propagation formula (or delta rule) [Haralick, 1994]. In the case of oriented artifacts, which introduce large errors to the gradients of the original pattern of fingerprint ridges, a different approach needs to be taken. In particular, the delta rule relies on the assumption that the first order Taylor approximation is valid.
The remaining method for which we would like to show the Fourier domain representation and its relationship with the methods seeking maximum energy orientation is the spatial window method.

In this method, an analysis window is positioned in a neighborhood. The window is then rotated and for each rotation of the window the 2-D signal present in the window is projected onto one of the window’s axes. The underlying principle is that the variance along the direction normal to the orientation of the ridge lines attains its maximum. Subsequent rotations of the spatial window and calculation of the variance of each projection result in an orientation signature. It is shown below that it is equivalent to the orientation signature obtained from the 2-D Fourier spectrum.

Before proceeding further, note that the rotation of the window in the spatial domain can be expressed in the Fourier domain as a translation in the angle variable $\alpha$. This approach can be applied to any pattern that can be represented as a superposition of harmonics. The spatial-frequency variables can be represented in a parametric form as

\begin{align}
  u &= r \cos(\alpha) \\
  v &= r \sin(\alpha)
\end{align}

(6-12)

The following derivations are for the fixed $\alpha$. In the summary we show that it can be extended to any $\alpha$. In the general case of the signal $s'(x,y)$ with Fourier Transform $S'(u,v)$, we have the following projection signature onto the x-axis:

\begin{align}
  s'(x) &= \int_{-\infty}^{\infty} s'(x,y)dy = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} S'(u,v)e^{jux}dudvdy = \\
  &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} S'(u,v)e^{-jux} \delta(v)dudv = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S'(u,0)e^{-jux} du
\end{align}

(6-13)

where we used the following definition,

\begin{align}
  \delta(v) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{jv x} dx
\end{align}

(6-14)

referred to as a model for the uncertainty principle in [Benedetto, 1997]. If the integration interval is finite, equation (6-14) leads to a sinc expression. The mean of the projection is
\[
\mu_{s_x} = \int_{-\infty}^{\infty} s'_x(x) dx = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} s''(u,0)e^{jux} du dx = \int_{-\infty}^{+\infty} s''(u,0) \delta(u) du = S''(0,0) \tag{6-15}
\]

and the variance for energy signals is

\[
\sigma^2_{s_x} = \int_{-\infty}^{\infty} s'_x(x) dx = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left| s''(u,0) e^{jux} du - S''(0,0) \right|^2 dx = \\
= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} s''(u',0) e^{jux} du' - S''(0,0) \left( \frac{1}{2\pi} \int_{-\infty}^{+\infty} s''(u,0) e^{-jux} du \right) dx = \\
= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} s''(u',0) s''(u,0) e^{jux} du' du - \frac{1}{2\pi} \int_{-\infty}^{+\infty} s''(u',0) S''(0,0) e^{jux} du' du - \\
\frac{1}{2\pi} \int_{-\infty}^{+\infty} s''(0,0) s''(u,0) e^{jux} du' du + \left( S''(0,0) \right)^2 dx 
\]

After applying (6-14), the result of (6-16) can be further simplified as

\[
\sigma^2_{s_x} = \frac{1}{2\pi} \left( \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} s''(u',0)(s''(u,0))' \delta(u'-u) du' du - \int_{-\infty}^{+\infty} s''(u',0) S''(0,0) \delta(u') du' - \\
\int_{-\infty}^{+\infty} s''(0,0)(s''(u,0))' \delta(u) du' du + \int_{-\infty}^{+\infty} s''(0,0)^2 dx \right) = \\
= \frac{1}{2\pi} \int_{-\infty}^{+\infty} s''(u,0)^2 du - 2 \int_{-\infty}^{+\infty} s''(0,0)^2 du + \int_{-\infty}^{+\infty} (s''(0,0))^2 dx 
\]

It is finite if the mean of \( s'(x,y) \) is zero, from which we get

\[
\sigma^2_{s_{x(0)}} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} s''(u,0)^2 du = \\
\frac{1}{2\pi} \left( \int_{0}^{+\infty} |s'(r \cos(\alpha) = 0, r \sin(\alpha) = 0)|^2 dr + \int_{0}^{+\infty} |s'(r \cos(\alpha = \pi), r \sin(\alpha = \pi))|^2 dr \right) \tag{6-18}
\]
From this equation one can obtain an orientation signature by observing that the rotation in polar coordinates is expressed as translation in \( \alpha \), and remembering that orientation is defined in \([-\pi/2, \pi/2]\), i.e. it is \( \pi \) periodic.

\[
\sigma^2_{x_0}(\alpha) = \frac{1}{2\pi} \left( \int_0^{+\infty} |S'(r \cos(\alpha), r \sin(\alpha))|^2 dr + \int_0^{+\infty} |S'(r \cos(\alpha + \pi), r \sin(\alpha + \pi))|^2 dr \right) \tag{6-19}
\]

Due to the property of symmetry of the power spectrum, \( \sigma^2_{x_0}(\alpha) \) can be written as

\[
\sigma^2_{y_0}(\alpha) = \frac{1}{\pi} \left( \int_0^{+\infty} |S'(r \cos(\alpha), r \sin(\alpha))|^2 dr \right) \tag{6-20}
\]

As can be seen from the equation (6-20) above, the orientation signature in the spatial window approach (the energy of the projection onto axis \( x \) of the rotated window) and the Fourier Transform approach can be related to each other using the formula that is essentially a form of a Parseval theorem. In the practical discrete implementation, the spatial window method is not very efficient; it requires rotation and interpolation of the pixel intensities. Equivalently, the orientation estimate can be found using the Fourier Transform. It can be accomplished, as shown in the previous chapter, using the Pseudo-Polar Fourier Transform. The algorithm for its fast computation is of the same complexity as the FFT, cf. [Keller et al., 2005] and references therein. Equation (6-20) represents the energy along the orientation \( \alpha \). If one considers the signal \( s'(x,y) \) in the Polar Fourier Transform representation \( S'(r,\alpha) \), then integration of the power spectrum over \( r \) for fixed \( \alpha \) and \( \alpha + \pi \) results in the equation (6-20). Orientation estimation using bandpass filters adds more robustness to the orientation estimation at the expense of increased computational complexity. It is presented in the following sections in the case of the additive noise and anisotropic defects.

### 6.2 Estimation in the Presence of Noise

The effect of the additive noise on the orientation estimation can be illustrated using a sine wave and Gaussian noise, according to the model proposed and presented in Chapter 3. Neither the sources of noise nor the mechanisms that govern the noise are studied in this work. However, one can make some assumptions. For instance, one can consider a quantization noise resulting
from using a finite binary precision. In usual cases, this type of noise can be expected to be negligibly low. The precision of 8 bits results in good quality of fingerprint images. Another type of noise to consider originates in the analogue part of acquisition device, namely, crosstalk between sensing cells or data lines in solid state sensors. The important factor that affects the quality in the fingerprint images the most originates in the interaction of the sensor and the user. The residue on the sensing surface and the skin properties introduce significant variability to the fingerprint images. Here we present results that can help explain how additive noise affects orientation estimates.

Suppose the signal has the following form

$$s(x, y) = A(0.5 + 0.5 \sin(w_0 x + v_0 y))$$  \hspace{1cm} (6-21)

and is corrupted by an additive IID zero-mean Gaussian noise with power spectral density (PSD) $P_N$. In order to find the dominant orientation of a pattern, one can proceed by designing a bank of filters $H_{\theta, r, b}(u, v)$ that are localized at given spatial frequencies, specified by the triplet of parameters. $\theta$ and $r$ jointly specify the spatial frequency of the center of the filter, and $b$ is the bandwidth. For instance, for the filter tuned to the signal $s(x,y)$ we would like $r = r_0 = \sqrt{u_0^2 + v_0^2}$ and $\theta = \angle(u_0, v_0)$. In practice, a finite number of filters is sufficient. We denote them as triplets

$$(\theta_k, r_q, b_r) \quad \text{where} \quad k = \{0,1,2...,K-1\}, \quad q = \{0,1,2...,Q-1\}, \quad r = \{0,1,2...,R-1\}$$

For the purpose of the orientation estimation in periodic patterns of known frequency, it is sufficient for the filters to have the $r$ parameter fixed, $r=r_0$. We will also assume the bandwidth $b$ is fixed. The dependence on $r_0$, $b_0$ can therefore be dropped. We will write $H_{\theta_k}(u, v)$. Each of the filters can be viewed as a bin. The energy captured in each bin can be used to indicate the presence of a signal. The maximum energy orientation is the orientation normal to the orientation of the ridges. The resolution of the filtering approach is limited by a theoretical bound determined by the uncertainty principle which relates the size of the window in the spatial domain to the size of the window in the spatial frequency domain.

The filtering in the spatial frequency domain can be accomplished via point-wise multiplication.
\[ S_k(u,v) = H_{\theta_k}(u,v)S(u,v) \quad (6-22) \]

and the power of the filtered signal equals

\[ P_{s,k} = \frac{1}{4\pi^2} \int\int_{-\infty}^{\infty} |H_{\theta_k}(u,v)S(u,v)|^2 \, du \, dv \quad (6-23) \]

Assuming that the signal is present within the pass band of one of the filters \( H_{\theta_k}(u,v) \) from the bank of filters, the power \( P_k \) reaches a maximum for certain values of the parameter \( \theta_k \).

Fourier Transform and oriented window methods rely on the energy calculated for the spatial frequencies along the radius from 0 to half of the sampling frequency. The orientation estimation in the presence of noise can be illustrated using the concept of an ideal filter, defined as

\[ H_{\theta_k}(u,v) = \begin{cases} 1 & \text{for } r_1 \leq \sqrt{u^2 + v^2} \leq r_2, \\
\angle(u,v) \in \left[ \theta_k - \frac{\Delta\theta}{2}, \theta_k + \frac{\Delta\theta}{2} \right] \cup \left[ \theta_k - \frac{\Delta\theta}{2} + \pi, \theta_k + \frac{\Delta\theta}{2} + \pi \right] \\
0 & \text{otherwise} \end{cases} \quad (6-24) \]

In the Fourier Transform method and spatial window methods of orientation estimation, we will assume the range of spatial frequencies is delimited by

\[ r_1 = 0, \quad r_2 = F_s / 2 \quad (6-25) \]

where \( F_s \) is the sampling frequency used in the discrete case. The orientation of the filter can be written as a function of the index \( k \) and the number of bins \( K \).

\[ \Delta\theta = \frac{\pi}{K}, \quad \theta_k = \frac{k\pi}{K} \quad \text{or} \quad \frac{k\pi}{2K} + \frac{\pi}{2K} \]

\[ \angle(u,v) \in \left[ \frac{k\pi}{K} - \frac{\pi}{2K}, \frac{k\pi}{K} + \frac{\pi}{2K} \right] \cup \left[ \frac{k\pi}{K} - \frac{\pi}{2K} + \pi, \frac{k\pi}{K} + \frac{\pi}{2K} + \pi \right] \]
The power spectrum of the signal \( s(x,y) \), excluding the DC component, which is insignificant since it contributes an offset to the orientation signature, can be written as

\[
|S(u,v)|^2 = \frac{A^2}{4} \left( \delta(u-u_0, v-v_0) + \delta(u+u_0, v+v_0) \right)
\]

(6-27)

The power of a 2-D sinusoidal signal with amplitude 0.5A (with zero DC component) is equal to

\[
P_s = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |S(u,v)|^2 \, du \, dv = \frac{1}{8\pi^2} A^2
\]

(6-28)

Assuming that the power spectral density (PSD) of the noise is known, we denote it as \( P_N \). The power of noise captured by each of the filters is denoted as \( P_{n,k} \). Calculated in the polar coordinate system, it can be written as

\[
P_{n,k} = \frac{1}{4\pi^2} \int_{\theta_k}^{\theta_k + \Delta \theta} \int_{r_1}^{r_2} P_N r^2 dr d\theta + \frac{1}{4\pi^2} \int_{\theta_k}^{\theta_k + \Delta \theta + \pi} \int_{r_1}^{r_2} P_N r^2 dr d\theta = \frac{1}{2\pi^2} \int_{\theta_k}^{\theta_k + \Delta \theta} \int_{r_1}^{r_2} P_N r^2 dr d\theta = \frac{\Delta \theta}{2\pi^2} P_N \left( \frac{r_2^2}{2} - r_1^2 \right)
\]

(6-29)

Substituting (6-25), (6-26) into equation (6-29) above results in

\[
P_{n,k} = \frac{F_s^2}{8\pi K} P_N = P_n
\]

(6-30)

In the case of signal detection and demodulation, it is sufficient to consider signal-to-noise ratio in order to determine the quality of the received signal. In the case of the orientation estimation, one needs to consider the signal-to-noise ratio in each of the K filters from the filter bank.

A probabilistic approach has been presented in the previous section. Here we limit our approach to the normalized orientation signature \( p_s(\theta) \):
\[ p_s(\hat{\theta}_k) = \frac{P_k}{\sum_{k=0}^{K-1} P_{n,k} + P_s} \]  

where \( P_k = P_{n,k} + P_{s,k} \) if the bin contains signal and \( P_k = P_{n,k} \) otherwise. Now, considering the case presented so far where all of the filters have the same bandwidth and transfer function magnitudes, and their Fourier Transforms do not overlap, for the bin that contains signal we have

\[ p_s(\hat{\theta}_k) = \frac{1 + P_s / P_n}{N + P_s / P_n} \]  

and otherwise, when only noise is present

\[ p_s(\hat{\theta}_{n,k}) = \frac{1}{N + P_s / P_n} \]  

where we used \( \hat{\theta}_{n,k} \) to denote the orientation estimate corrupted by noise. It is more informative to examine the SNR, which can be used as a proxy for how likely it is that a given bin \( k \) contains the signal. From (6-32), and (6-33) we have

\[ \frac{p_s(\hat{\theta}_k)}{p_s(\hat{\theta}_{n,k})} = 1 + \frac{P_s}{P_n} \]  

After using (6-28), (6-30) in (6-34) we obtain

\[ \frac{p_s(\hat{\theta}_k)}{p_s(\hat{\theta}_{n,k})} = 1 + \frac{A^2 K}{\pi f_s^2 P_n} \]  

which shows as a function of the number of bins \( K \), sampling frequency and noise spectral density, how much more likely we are to find the true orientation than not. In the case when the filters do not have disjoint frequency responses, the signal is present in multiple bins. This case can be illustrated using the Gabor filter bank approach to orientation estimation presented in the previous section. Because of our assumption of the constant noise PSD and equality of the filters.
up to the rotation parameter, the amount of noise present in each bin is the same. However, because of the infinite support of Gabor filters, a small fraction of the signal is present in every bin. This can be written as a function of the distance between the center of the filter’s spatial bandwidth and the signal. The frequency response of the Gabor filter has the form of a pair of Gaussians $H_{1,\theta_k}, H_{2,\theta_k}$, denoted as $H_{\theta_k} = H_{1,\theta_k} + H_{2,\theta_k}$, where $H_{1,\theta_k}, H_{2,\theta_k}$ are identical images of one another symmetric about the origin. We can expect that the bandwidth of the Gabor filter is small in comparison to the spatial frequency of the signal $r_0$. We can also make an assumption that the center frequency of the filter coincides with the signal at exactly one of the orientations $\theta_k$.

In order to make the calculations simple, we assume without loss of generality that

$$r_0 = |v_0|, \quad u_0 = 0 \quad (6-36)$$

It corresponds to the pattern of ridges parallel to the x-axis. The frequency response of the Gabor filter is equal to

$$H_{\theta_k}(u,v) = H_{1,\theta_k} + H_{2,\theta_k} = 2\pi \sigma^2 e^{-\frac{(u-u_k)^2 + (v-v_k)^2}{2}} + 2\pi \sigma^2 e^{-\frac{(u+u_k)^2 + (v+v_k)^2}{2}} \quad (6-37)$$

$$= 2\pi \sigma^2 \left( e^{-\frac{(u-u_k)^2 + (v-v_k)^2}{2}} + e^{-\frac{(u+u_k)^2 + (v+v_k)^2}{2}} \right)$$

$$u = r \cos(\theta) \quad v = r \sin(\theta)$$

$$u_k = r_0 \cos(\theta_k) \quad v_k = r_0 \sin(\theta_k)$$

The filter in the spatial domain is represented as a product of a cosine and a circularly symmetric Gaussian.

$r_0$ is used to denote the fact that we assume that the filters are tuned to the spatial frequency of the signal. We will also need $|H_{\theta_k}(u,v)|^2$ in order to find the power spectrum of the filtered signal.
\begin{align*}
|H_{\delta_k}(u,v)|^2 &= 4\pi^2 \sigma^4 \left( e^{-\sigma^2((u-u_0)^2+(v-v_0)^2)} + 2e^{-\sigma^2(u_0^2+v_0^2)} + e^{-\sigma^2((u+u_0)^2+(v+v_0)^2)} \right) \\
&= 4\pi^2 \sigma^4 \left( e^{-\sigma^2((u-u_0)^2+(v-v_0)^2)} + 2e^{-\sigma^2(u^2+v^2)} + e^{-\sigma^2((u+u_0)^2+(v+v_0)^2)} \right) \\
\text{and the volume under the curve above (sum of Gaussians)}
\end{align*}

\begin{align*}
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |H_{\delta_k}(u,v)|^2 \, du \, dv &= 4\pi^2 \sigma^4 \left( \frac{\pi}{\sigma^2} + 2e^{-\sigma^2 r_0^2} \frac{\pi}{\sigma^2} + \frac{\pi}{\sigma^2} \right) = 8\pi^3 \sigma^2 \left( 1 + e^{-\sigma^2 r_0^2} \right)
\end{align*}

The power of the filtered signal is

\begin{align*}
P_{x,k} &= \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |S(u,v)|^2 |H_{\delta_k}(u,v)|^2 \, du \, dv = \frac{1}{4\pi^2} A^2 w_k^2 = \frac{A^2}{2} \sigma^4 \left( e^{-\sigma^2((u+u_0)^2+(v+v_0)^2)} + e^{-\sigma^2((u-u_0)^2+(v-v_0)^2)} + 2e^{-\sigma^2 r_0^2} e^{-\sigma^2 r_0^2} \right) = \frac{A^2}{2} \sigma^4 e^{-2\sigma^2 r_0^2} \left( e^{2\sigma^2 r_0^2} \sin(\theta_k) + e^{-2\sigma^2 r_0^2} \sin(\theta_k) + 2e^{\sigma^2 r_0^2} e^{-\sigma^2 r_0^2} \cos^2(\theta_k) \right)
\end{align*}

The weight \( w_k \) is defined below and will be useful in the next section

\begin{align*}
w_k^2 &= \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(u-u_0,v-v_0) + \delta(u+u_0,v+v_0) \left| H_{\delta_k}(u,v) \right|^2 \, du \, dv = 8\pi^2 \sigma^4 e^{-2\sigma^2 r_0^2} \left( e^{2\sigma^2 r_0^2} \sin(\theta_k) + e^{-2\sigma^2 r_0^2} \sin(\theta_k) + 2e^{\sigma^2 r_0^2} e^{-\sigma^2 r_0^2} \cos^2(\theta_k) \right)
\end{align*}

The power of noise captured by each filter \( H_{\delta_k}(u,v) \) is equal to the PSD \( P_N \) times the volume under \( |H_{\delta_k}(u,v)|^2 \) (6-39), as shown below.

\begin{align*}
P_n &= \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P_N |H_{\delta_k}(u,v)|^2 \, du \, dv = \frac{P_N}{4\pi^2} 8\pi^3 \sigma^2 \left( 1 + e^{-\sigma^2 r_0^2} \right) = 2P_N \pi \sigma^2 \left( 1 + e^{-\sigma^2 r_0^2} \right)
\end{align*}
It is now possible to calculate the ratio (6-34) for the Gabor filtering approach using (6-40) and (6-42).

\[
\frac{p_s(\hat{\theta}_k)}{p_s(\hat{\theta}_{n,k})} = 1 + \frac{P_s}{P_n} = 1 + \frac{A^2 \sigma^2 e^{-2\sigma^2 r_0^2}}{4PN\pi (1 + e^{-2\sigma^2 r_0^2})} \left( e^{2\sigma^2 r_0^2 \sin(\theta_k)} + e^{-2\sigma^2 r_0^2 \sin(\theta_k)} + 2e^{\sigma^2 r_0^2} e^{-\sigma^2 r_0^2 \cos^2(\theta_k)} \right)
\]

The ratio will reach a maximum when the signal is located at the center of \(H_{1,\theta_k}\) or \(H_{2,\theta_k}\). Here it is attained for \(\theta_k = \pi/2\) as shown below

\[
\frac{p_s(\hat{\theta}_k)}{p_s(\hat{\theta}_{n,k})} = 1 + \frac{A^2 \sigma^2}{4PN\pi (1 + e^{-2\sigma^2 r_0^2})} \left( 1 + e^{-4\sigma^2 r_0^2} + 2e^{-\sigma^2 r_0^2} \right) \approx 1 + \frac{A^2 \sigma^2}{4PN\pi}
\]

The simplification above holds when the product \(\sigma r_0\) is large

\[e^{-\sigma^2 r_0^2} < 0.1 \iff \sigma r_0 > \ln(10) \approx 2.30\]

A comparison of (6-35) and the simplified form on the right hand side of (6-44) allows one to decide which method might outperform the other and for what values of the filter parameters. It does so. Subtracting (6-35) from (6-44) gives

\[
SNR_{dif} = \frac{A^2}{PN} \left( \frac{\sigma^2}{4\pi} - \frac{K}{\pi P_s^2} \right).
\]

\(SNR_{dif}\) is positive when the Gabor filter bank approach shows better performance. In the Fourier Transform method the SNR decreases as a function of the sampling frequency. Bandpass filtering is necessary to remove this dependence.
6.3 *Estimation in the Presence of Anisotropic Defects*

Oriented defects present more difficulty in the analysis of their effects on the orientation estimate. They are modeled as a multiplicative mask. The Fourier Transform of the signal distorted by such a mask is given in Chapter 3 and is repeated below. In the following part the more concise matrix notation is used.

\[
S'(u) = S(u) + M(u) - M(u) * S(u)
\]  \hspace{1cm} (6-46)

\[
u = [u \quad v]^T
\]

The orientation estimate is found from an orientation signature and, as shown so far, it can be obtained from the power spectrum. The case presented below is simplified for clarity. Similarly as in the previous section, we use the signal \( s(x) \) in the form of a sine wave given by (6-21). The oriented defect is modeled as a Gaussian function.

\[
m(x) = B \exp \left( \frac{-(x-x_0)^T \Sigma_o^{-1} (x-x_0)}{2} \right)
\]  \hspace{1cm} (6-47)

\[
x = [x \quad y]^T
\]

\[
x_o = [x_o \quad y_o]^T
\]

\( \Sigma_o \) is the covariance matrix, \( B=1 \) in the defect model (3–5). We will need the Fourier Transform of \( m(x) \):

\[
M(u) = 2\pi B^{1/2} \epsilon \Sigma_o^{1/2} \exp \left( -\frac{u^T \Sigma_o^{-1} u}{2} \right) \exp \left( -ju^T x_o \right).
\]  \hspace{1cm} (6-48)

For the sake of clarity \( x_o=(0,0) \); the imaginary part of the equation above is then 0. The power spectrum of the signal can be computed as follows

\[
\hat{P}_x(u) = |S'(u)|^2 = \text{Re}\{S'(u)\}^2 + \text{Im}\{S'(u)\}^2
\]  \hspace{1cm} (6-49)
The Fourier Transform of the signal \( s'(x) \) – a sine wave multiplied by a Gaussian used as a model of an oriented defect, is given by

\[
S'(u) = \frac{A}{2} \delta(u) - j \frac{A}{4} \left( \delta(u - u_o) - \delta(u + u_o) \right) + Ce^{\left( -\frac{u^T \Sigma_o u}{2} \right)} \frac{AC}{2} e^{\left( -\frac{u^T \Sigma_o u}{2} \right)} +
\]

\[
j \frac{AC}{4} \left\{ \begin{align*}
&\frac{\left( -(u-u_o)^T \Sigma_o (u-u_o) \right)}{2} - \frac{\left( -(u+u_o)^T \Sigma_o (u+u_o) \right)}{2} \\
&\frac{\left( -(u-u_o)^T \Sigma_o (u-u_o) \right)}{2} - \frac{\left( -(u+u_o)^T \Sigma_o (u+u_o) \right)}{2}
\end{align*} \right\}
\]

\[
u_o = [\nu_o \ v_o]^T
\]

The following definition has been introduced

\[
C = 2\pi B |\Sigma_o|^{1/2}
\]

The equation (6-49) can be now calculated by finding the square of real and imaginary parts in (6-50) as shown below. In the case of the real part we have

\[
\text{Re}\{S'(u)\} = \frac{A}{2} \delta(u) + C \left( 1 - \frac{A}{2} \right) e^{-u^T \Sigma_o u}
\]

\[
\text{Re}\{S'(u)\}^2 = \left( \frac{A^2}{4} + AC \left( 1 - \frac{A}{2} \right) \right) \delta(u) + C^2 \left( 1 - \frac{A}{2} \right)^2 e^{-2u^T \Sigma_o u}
\]

Similarly, for the imaginary part

\[
\text{Im}\{S'(u)\} = \frac{A}{4} \left\{ \begin{align*}
&\delta(u - u_o) - \delta(u + u_o) - C \left( e^{\left( -\frac{(u-u_o)^T \Sigma_o (u-u_o)}{2} \right)} - e^{\left( -\frac{(u+u_o)^T \Sigma_o (u+u_o)}{2} \right)} \right) \\
&\delta(u - u_o) - \delta(u + u_o) - C \left( e^{\left( -\frac{(u-u_o)^T \Sigma_o (u-u_o)}{2} \right)} - e^{\left( -\frac{(u+u_o)^T \Sigma_o (u+u_o)}{2} \right)} \right)
\end{align*} \right\}
\]

126
\[\text{Im}[S'(u,v)]^2 = \left(-\frac{A}{4}\right)^2 \left[\delta\left(u-v, u_0\right) - \delta\left(u + v, u_0\right) - C \left(e^{-\frac{(u-v)^2}{2}} - e^{-\frac{(u+v)^2}{2}}\right)\right]^2 = \]

\[= \frac{A^2}{16} \left[\delta\left(u-v, u_0\right) + \delta\left(u + v, u_0\right) - 2C\left[\delta\left(u-v, u_0\right) - \delta\left(u + v, u_0\right)\right] \left(e^{-\frac{(u-v)^2}{2}} - e^{-\frac{(u+v)^2}{2}}\right)\right] = \]

\[A^2C^2 \left[\left(-e^{2(u-v)^2} + 2\right)\left[\delta\left(u-v, u_0\right) + \delta\left(u + v, u_0\right)\right] = \frac{A^2}{16} \left[\delta\left(u-v, u_0\right) + \delta\left(u + v, u_0\right)\right] + \right.\]

\[\left.\frac{A^2C^2}{8} \left[1 + C\left(e^{-2(u-v)^2} + e^{-2(u+v)^2}\right)\left[\delta\left(u-v, u_0\right) + \delta\left(u + v, u_0\right)\right]\right] = \right.\]

\[= A^2 \left[\frac{1}{16} + \frac{C}{4}\right] \left[\delta\left(u-v, u_0\right) + \delta\left(u + v, u_0\right)\right] + \]

\[\frac{A^2C^2}{16} \left[\left(-2e^{2(u-v)^2} - 2e^{2(u+v)^2}\right)\left[\delta\left(u-v, u_0\right) + \delta\left(u + v, u_0\right)\right]\right] = \]

\[\text{(6-55)}\]

The power spectrum (6-49) can now be rewritten using the results for squared real and imaginary parts, as shown below

\[P_u(u) = \left(\frac{A^2}{4} + AC\left(1 - \frac{A}{2}\right)\right)\delta(u) + C^2\left(1 - \frac{A}{2}\right) e^{-u^2} + \]

\[A^2 \left[\frac{1}{16} + \frac{C}{4}\right] \left[\delta(u-v, u_0) + \delta(u + v, u_0)\right] + \]

\[\frac{A^2C^2}{16} \left[\left(-2e^{2(u-v)^2} - 2e^{2(u+v)^2}\right)\left[\delta(u-v, u_0) + \delta(u + v, u_0)\right]\right] = \]

\[\text{(6-56)}\]

Equation (6-56) can be simplified when \(e^{-u^2}\) is a small constant. This is a close approximation in the case in which we are interested, i.e. the anisotropic defect crosses the ridge pattern at 90 degrees.
Equation (6-57) can be used to illustrate the orientation estimation. Let us start with the Gabor filtering method; the Gabor filter is defined in (6-37), but here it is shown using the matrix notation.

\[
\begin{align*}
P_r(u) &\approx \left(\frac{A^2}{4} + AC\left(1 - \frac{A}{2}\right)\right)\delta(u) + C^2\left(1 - \frac{A}{2}\right)^2 \cdot e^{-w^T \Sigma u} + \\
A^2\left(\frac{1}{16} - \frac{C}{4}\right)(\delta(u - u_o) + \delta(u + u_o)) + \frac{A^2C^2}{16}\left(e^{-(u-u_o)^T \Sigma (u-u_o)} + e^{-(u+u_o)^T \Sigma (u+u_o)}\right)
\end{align*}
\]

Equation (6-57) can be used to illustrate the orientation estimation. Let us start with the Gabor filtering method; the Gabor filter is defined in (6-37), but here it is shown using the matrix notation.

\[
H_{\theta_k}(u) = 2\pi|\Sigma_k|^{1/2} \left(\begin{array}{c}
\frac{-(u-u_k)^T \Sigma_k (u-u_k)}{2} e^{-\frac{(u-u_k)^T \Sigma_k (u-u_k)}{2}} + e^{-\frac{(u+u_k)^T \Sigma_k (u+u_k)}{2}}
\end{array}\right)
\]

For the purpose of evaluation of the orientation signature, the DC component of the signal (the term containing \(\delta(u)\)) will be omitted. The modified power spectrum \(P'\) can be then written as

\[
P'_r(u) = C^2\left(1 - \frac{A}{2}\right)^2 \cdot e^{-w^T \Sigma u} + A^2\left(\frac{1}{16} - \frac{C}{4}\right)(\delta(u - u_o) + \delta(u + u_o)) + \\
\frac{A^2C^2}{16}\left(e^{-(u-u_o)^T \Sigma (u-u_o)} + e^{-(u+u_o)^T \Sigma (u+u_o)}\right)
\]

We will use \(w_k\), defined in the equation (6-41) in the previous section, and the fact that the product of two Gaussians is another Gaussian. The method presented below is applicable to the case of any spatial frequency of the signal sine wave and to any defect modeled as a Gaussian mask. Here, for clarity, it is assumed that the sine wave's spatial frequency is \(u_o=[0 v_o]^T\) as in the previous section where we looked at the case of additive noise.

\[
P_{s,k} = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P'_r(u)\left|H_{\theta_k}(u)\right|^2 du
\]

The squared modulus of the filter frequency response is given by equation (6-61)
\[ |H_{\theta_k}(u)|^2 = 4\pi^2|\Sigma_k| \left( e^{-(u-u_k)^T \Sigma_k (u-u_k)} + 2e^{-(u-u_k)^T \Sigma_k u_k e^{-u^T \Sigma_k u}} + e^{-(u+u_k)^T \Sigma_k (u+u_k)} \right) \quad (6-61) \]

\[ |\Sigma_k| = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix} = \sigma^4 \]

It can be simplified by noting that the spatial frequency bandwidth of the Gabor filter is narrow, i.e. \(1/\sigma \ll \nu_0\), and therefore the second term in the parentheses in the equation (6-61) above can be omitted. Then we have

\[ |H_{\theta_k}(u)|^2 \approx 4\pi^2|\Sigma_k| \left( e^{-(u-u_k)^T \Sigma_k (u-u_k)} + e^{-(u+u_k)^T \Sigma_k (u+u_k)} \right) \quad (6-62) \]

We are interested in the power that is captured by the k-th filter; it can be calculated for each term in (6-59) separately. Using the results from the previous section (6-41), we can observe that the term in (6-59) that contains delta distributions contributes power equal to

\[ \int \int A^2 \left( \frac{1}{16} - \frac{C}{4} \right) (\delta(u - u_o) + \delta(u + u_o)) \cdot |H_{\theta_k}(u)|^2 \, du = \frac{1}{4\pi^2} A^2 \left( \frac{1}{16} - \frac{C}{4} \right) \omega_k^2 = \]

\[ = \frac{1}{4\pi^2} A^2 \left( \frac{1}{16} - \frac{C}{4} \right) 8\pi^2 \sigma^4 e^{-2\sigma^2 r_o^2} \left( e^{2e^{2r_o^2 \sin(\theta_o)}} + e^{-2e^{2r_o^2 \sin(\theta_o)}} + 2e \sigma^2 e^{-2r_o^2 \cos^2(\theta_o)} \right) \]

\[ = A^2 \left( \frac{1}{8} - \frac{C}{2} \right) \sigma^4 e^{-2\sigma^2 v_o^2} \left( e^{2\sigma^2 v_o^2 \sin(\theta_o)} + e^{-2\sigma^2 v_o^2 \sin(\theta_o)} + 2e \sigma^2 e^{-\sigma^2 v_o^2 \cos^2(\theta_o)} \right) \quad (6-63) \]

The remaining terms in (6-59) contain Gaussian functions. Their product with the squared modulus of the transfer function of Gabor filter is a sum of Gaussians. Consequently the evaluation of equation (6-60) is simplified. Below we show that the product of any two Gaussians is a Gaussian. The product of two Gaussians can be written as

\[ e^{-(u-u_k)^T \Sigma_k (u-u_k)} - (u-u_o)^T \Sigma_o (u-u_o) \]

We need to show that the sum of quadratics in the exponent is equal to
\( \mathbf{u} \mathbf{u}^T \mathbf{\Sigma}_k \mathbf{u} + \mathbf{u} \mathbf{u}^T \mathbf{\Sigma}_o \mathbf{u} = \mathbf{u} \mathbf{u}^T + \text{const} \)

Completing of the square on the left hand side gives

\[ \mathbf{u}^T \left( \mathbf{\Sigma}_k + \mathbf{\Sigma}_o \right) \mathbf{u} - 2 \mathbf{u}^T \left( \mathbf{\Sigma}_k \mathbf{u}_k + \mathbf{\Sigma}_o \mathbf{u}_o \right) + \mathbf{u}_k^T \mathbf{\Sigma}_k \mathbf{u}_k + \mathbf{u}_o^T \mathbf{\Sigma}_o \mathbf{u}_o \]

We can multiply the second term of the quadratic above by the identity matrix \( \mathbf{I} = \mathbf{\Lambda} \mathbf{\Lambda}^{-1} \), where \( \mathbf{\Lambda} \) is the covariance matrix of the product and is equal to

\[ \mathbf{\Lambda} = \mathbf{\Sigma}_k + \mathbf{\Sigma}_o , \]

which leads to

\[ \mathbf{u}^T \mathbf{\Lambda} \mathbf{u} - 2 \mathbf{u}^T \mathbf{\Lambda} \mathbf{\Lambda}^{-1} \left( \mathbf{\Sigma}_k \mathbf{u}_k + \mathbf{\Sigma}_o \mathbf{u}_o \right) + \mathbf{u}_k^T \mathbf{\Sigma}_k \mathbf{u}_k + \mathbf{u}_o^T \mathbf{\Sigma}_o \mathbf{u}_o \]

(6-64)

Define the mean of the product of Gaussians as

\[ \mathbf{\mu} = \mathbf{\Lambda}^{-1} \left( \mathbf{\Sigma}_k \mathbf{u}_k + \mathbf{\Sigma}_o \mathbf{u}_o \right) \]

(6-65)

We can now add and subtract the term

\[ \mathbf{\mu}^T \mathbf{\Lambda} \mathbf{\mu} \]

to/ from (6-64) as it does not change the value of the expression. The resulting expression contains a quadratic form and a constant as sought.

\[ \mathbf{u}^T \mathbf{\Lambda} \mathbf{u} - 2 \mathbf{u}^T \mathbf{\Lambda} \mathbf{\mu} + \mathbf{\mu}^T \mathbf{\Lambda} \mathbf{\mu} - \mathbf{\mu}^T \mathbf{\Lambda} \mathbf{\mu} + \mathbf{u}_k^T \mathbf{\Sigma}_k \mathbf{u}_k + \mathbf{u}_o^T \mathbf{\Sigma}_o \mathbf{u}_o = \]

\[ = (\mathbf{u} - \mathbf{\mu})^T \mathbf{\Lambda} (\mathbf{u} - \mathbf{\mu}) - \mathbf{\mu}^T \mathbf{\Lambda} \mathbf{\mu} + \mathbf{u}_k^T \mathbf{\Sigma}_k \mathbf{u}_k + \mathbf{u}_o^T \mathbf{\Sigma}_o \mathbf{u}_o \]

Now, the product of two Gaussians can be written as

\[ e^{-\left( \mathbf{u} - \mathbf{\mu} \right)^T \mathbf{\Sigma}_k \left( \mathbf{u} - \mathbf{\mu} \right) - \left( \mathbf{u} - \mathbf{\mu} \right)^T \mathbf{\Sigma}_o \left( \mathbf{u} - \mathbf{\mu} \right)} = \]

\[ = e^{-\left( \mathbf{u} - \mathbf{\mu} \right)^T \mathbf{\Lambda} (\mathbf{u} - \mathbf{\mu}) - \mathbf{\mu}^T \mathbf{\Lambda} \mathbf{\mu} + \mathbf{u}_k^T \mathbf{\Sigma}_k \mathbf{u}_k - \mathbf{u}_o^T \mathbf{\Sigma}_o \mathbf{u}_o} \]

130

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
We will need the volume of the product of Gaussians shown above on the right hand side. It equals

\[
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(u-u_\mu)^T \Lambda (u-u_\mu)} e^{-(u-u_o)^T \Sigma_o} du = \pi |\Lambda|^{-1/2} e^{u_o^T \Lambda u_o} = \pi |\Lambda|^{-1/2} e^{u_o^T \Lambda u_o}.
\]

(6-67)

We can now complete the calculation of the power of the signal filtered using the Gabor filter. The first term in the equation (6-59) contributes power equal to

\[
\frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left| H_{\theta_k}(u) \right|^2 du =
\]

\[
\begin{align*}
&= \frac{1}{4\pi^2} \left| \Sigma_k \right| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(u-u_k)^T \Sigma_k (u-u_k)} + e^{-(u+u_k)^T \Sigma_k (u+u_k)} \left| \left| \left| \begin{array}{c}
\begin{array}{c}
0 \quad 0
\end{array}
\end{array}\right| \right|^2 du =
\end{align*}
\]

\[
= \frac{1}{4\pi^2} \left| \Sigma_k \right| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(u-u_k)^T \Sigma_k (u-u_k)} + e^{-(u+u_k)^T \Sigma_k (u+u_k)} \left| \left| \left| \begin{array}{c}
\begin{array}{c}
0 \quad 0
\end{array}
\end{array}\right| \right|^2 du =
\]

\[
= \left| \Sigma_k \right| \left| \left| \left| \begin{array}{c}
\begin{array}{c}
0 \quad 0
\end{array}
\end{array}\right| \right|^2 \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(u-u_k)^T \Sigma_k (u-u_k)} + e^{-(u+u_k)^T \Sigma_k (u+u_k)} \left| \left| \left| \begin{array}{c}
\begin{array}{c}
0 \quad 0
\end{array}
\end{array}\right| \right|^2 du =
\]

\[
= 2\pi C^2 \left| \Sigma_k \right| \left| \left| \left| \begin{array}{c}
\begin{array}{c}
0 \quad 0
\end{array}
\end{array}\right| \right|^2 \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(u-u_k)^T \Sigma_k (u-u_k)} + e^{-(u+u_k)^T \Sigma_k (u+u_k)} \left| \left| \left| \begin{array}{c}
\begin{array}{c}
0 \quad 0
\end{array}
\end{array}\right| \right|^2 du =
\]

\[
= 2\pi C^2 \left| \Sigma_k + \Sigma_o \right|^{-1/2} \left| \Sigma_k \right| e^{(\Sigma_k + \Sigma_o)^{-1} \Sigma_k u_k} = 2\pi C^2 \left| \Sigma_k + \Sigma_o \right|^{-1/2} \left| \Sigma_k \right| e^{(\Sigma_k + \Sigma_o)^{-1} \Sigma_k u_k}.
\]

(6-68)

The last line follows after substituting \( \Lambda \) and mean \( \mu \). \( \mu \) can be found from (6-65) after setting \( u_o = (0,0)^T \) as shown below

\[
\mu = (\Sigma_k + \Sigma_o)^{-1} \Sigma_k u_k
\]

Similarly, we can obtain the power contributed by the last term in (6-59)
\[ \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{A^2 C^2}{16} \left( e^{-(u-u_o)^T \Sigma_o (u-u_o)} + e^{-(u+u_o)^T \Sigma_o (u+u_o)} \right) \cdot |H_{\theta_k}(u)|^2 \, du = \]

\[ = \frac{1}{4\pi^2} A^2 C^2 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} 4\pi^2 |\Sigma_k| \left( e^{-(u-u_k)^T \Sigma_k (u-u_k)} + e^{-(u+u_k)^T \Sigma_k (u+u_k)} \right) \, du \]

\[ = 2 \frac{A^2 C^2}{16} |\Sigma_k| \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left( e^{-(u-u_k)^T \Sigma_k (u-u_k)} + e^{-(u+u_k)^T \Sigma_k (u+u_k)} \right) \, du = \]

\[ = \frac{A^2 C^2}{8} \left( \pi |\Lambda|^{-1/2} e^{\mu_1^T \Lambda \mu_1 - u_k^T \sigma_k^o} - u_k^T \sigma_k^o \right) \]

\[ \left( e^{(\Sigma_k + \Sigma_o)^{-1} (\Sigma_k u_k + \Sigma_o u_o)} + e^{(\Sigma_k + \Sigma_o)^{-1} (-\Sigma_k u_k + \Sigma_o u_o)} \right) \]

\[ (6-69) \]

In the derivation above we used the symmetry of the filter and the signal. The last line of the equation above follows after substituting for \( \Lambda \), the covariance matrix of the product of Gaussians and the two mean vectors

\[ \mu_1 = (\Sigma_k + \Sigma_o)^{-1} (\Sigma_k u_k + \Sigma_o u_o), \]

\[ \mu_2 = (\Sigma_k + \Sigma_o)^{-1} (-\Sigma_k u_k + \Sigma_o u_o) \]

of the product of Gaussians. Combining the results (6-63), (6-68), and (6-69) into one equation gives the power captured by the \( k \)-th rotation of the filter

\[ (6-70) \]

\[ P_{g,k} = A^2 \left( \frac{1}{8} - \frac{C}{2} \right) \sigma^4 e^{-2\sigma^2 v_o^2} \left( e^{2\sigma^2 v_o^2 \sin(\theta_k)} + e^{-2\sigma^2 v_o^2 \sin(\theta_k)} + 2e^{\sigma^2 v_o^2} e^{-\sigma^2 v_o^2 \cos^2(\theta_k)} \right) + \]

\[ + \pi C^2 |\Sigma_k + \Sigma_o|^{-1/2} |\Sigma_k| e^{-u_k^T \Sigma_k u_k} \cdot (P_{g,k} + P_{o,k}) \]

\( P_{g,k} \) and \( P_{o,k} \) are given by (6-72) and (6-73) respectively. The evaluation of the equation above can be rather tedious. For this reason, it is illustrated on a diagram. Let us consider closely the signal consisting of horizontal ridges i.e. \( u_o = 0 \). The defect is taken to be a Gaussian with an
arbitrary shape matrix $\Sigma_o$. It needs to be ensured that the simplifying assumptions: (6-56), (6-57) and (6-62) hold. In the following example we will assume that the eigenvalues are constant, but we will allow the rotation transformation to be applied to $\Sigma_o$.

$$\Sigma_o = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix}$$

$\sigma_x >> \sigma_y$ and $u_o = 0$

$$\Sigma_o' = R \Sigma_o R^T$$

$$R = \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix}$$

As it can be seen, there are two extreme situations: 1) when $\phi = 0$, the defect has the most severe effect on the orientation estimation through the low frequency spectral component, and 2) when $\phi = \pi/2$ the defect has no degrading effect on the orientation estimate. Figure 6.2 below depicts the orientation signature computed using (6-70) for the following three cases: $\phi = 0$ degrees, $\sigma_x = 1$, $\sigma_y = 7$ (first row); $\phi = -10$ degrees, $\sigma_x = 1$, $\sigma_y = 7$ (second row); and $\phi = -10$, $\sigma_x = .5$, $\sigma_y = 7$ (the last row). It appears that the two terms in the last line in (6-70) are significant. The first term,

$$P_{q,k} = 2 \left(1 - \frac{A^2}{2}\right) e^{\left(\Sigma_k + \Sigma_o\right)^{-1} \Sigma_k u_k} \Sigma_k u_k$$

(6-72)

carries the energy associated with the defect spectral component centered at spatial frequency (0,0) and is shown in the second column in the Figure 6.2. The second term

$$P_{o,k} = \frac{A^2}{8} e^{-u_o^T \Sigma_o u_o} \left(e^{\left(\Sigma_k + \Sigma_o\right)^{-1} \left(\Sigma_k u_k + \Sigma_o u_o\right)} \left(\Sigma_k u_k + \Sigma_o u_o\right) + e^{\left(\Sigma_k + \Sigma_o\right)^{-1} \left(-\Sigma_k u_k + \Sigma_o u_o\right)} \left(-\Sigma_k u_k + \Sigma_o u_o\right)} \right)$$

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)

(6-73)
Figure 6.2. The distribution of power in a function of the orientation signature for the three cases of defect orientation and shape; $\phi=0$, $\sigma_x=1$, $\sigma_y=7$ (first row), $\phi=-10$, $\sigma_x=1$, $\sigma_y=7$ (second row) and $\phi=-10$, $\sigma_x=5$, $\sigma_y=7$ (the last row). Orientation signature is obtained by translation of x-axis by 90 degrees.

In the case when the orientation is estimated using the power spectrum (PPFT or the spatial window methods), the orientation signature will contain the entire lowpass component due to the oriented defect. The approach shown in the previous section involves ideal filters. Following this approach in an attempt to demonstrate the behavior of the orientation signature leads to integrals that involve error functions. In order to obtain a closed form solution, analogous to (6-70), one can instead use a lowpass filter that has a form of a Gaussian function. Then the same approach as shown in this section for the case of Gabor filters can be used.
From the equations presented above, it can be seen that an oriented defect creates a low pass spectral component that needs to be filtered out in order to improve the robustness of the orientation estimate. Since for fingerprint images the signal associated with the ridge valley cycle is constrained to a relatively narrow spatial frequency band, it is possible to improve the orientation estimate in the presence of oriented defects via bandpass filtering. One could also make an attempt to estimate the frequency for each individual fingerprint and design an appropriate matched filter. Bandpass filtering, where the filters are matched to the local properties of the image, are referred to as contextual filtering. They have been presented in the literature and showed good performance in enhancement, ridge line orientation estimation and quality assessment of the fingerprint images. The most widely known example is the Gabor filter bank approach. Unfortunately, as the illustration above shows, the bandpass filtering does not allow for separating the signal from multiplicative defects.

### 6.4 Finite Window and Sampling Effects

The mathematical derivations shown in this chapter apply to continuous spatial and spatial frequency domains. The similar observations can be made for the discrete case. The methods are based on the principle of finding the orientation of maximum energy or MSE-optimal orientation. The principle of the conservation of energy holds for both continuous and discrete domains. The phenomena present in the discrete and finite dimensional case are: 1) processing gain – signal-to-noise ratio improves when the size of the analysis window increases; 2) leakage – whenever the ratio of the window length to the signal period is not an integer, the signal’s energy appears in all DFT bins, 3) so called “scalloping effect” [Lyons, 1997]. The equivalence of the two methods, spatial window and Fourier Transform based approach has been shown without considering the specific implementation. For example, in the spatial window method the rotation of the window requires interpolation in the discrete case, and the Fourier Transform method is implemented using PPFT algorithm.

The processing gain is related to a number of samples taken within the analysis window. This effect is rather well studied in the digital signal processing community. It is a useful property. If the spatial extent of the region, in which the periodicity of the signal is constant, is large, then even significant noise and local distortions become insignificant and the true orientation can be easily found from the spatial frequency location of the maximum of the power spectrum. In practice, interesting information is hidden in local changes of the signal and more localized analysis is required. The advantage of the processing gain is therefore limited.
The effect of signal energy leakage is directly related to the window size. The energy of the infinite duration sine wave is infinite and its frequency localization is described as the delta distribution. As we are forced to use a finite analysis window, the delta is replaced by the sinc function. It is an infinite duration signal with finite variance. That is, we can say that the energy is concentrated in the finite interval in the frequency domain. It is a known fact that the product of two signals in the spatial domain corresponds to convolution in the spatial frequency domain. Due to infinite length of the sinc function and because of the linearity of the convolution, the signal is present in every frequency bin. The exception is when the signal is periodic and its frequency components are such that the integer number of periods fits within the analysis window. Different windowing functions have been designed in order to decrease the leakage effect. However, this is always a tradeoff. The fundamental limit of what can be achieved through windowing is determined by the uncertainty principle.

The windowing effect is also attributed to the scalloping loss. This is related to the shape of the main lobe of the Fourier Transform of the window function. It means that the approximation of the energy achieves its maximum at the bin center and decreases towards the bin's boundary. It can be observed, for instance, by performing a DFT of a sine wave in a window that captures a non-integer number of signal's cycles. The signal will then be present in all bins and the main lobe will be represented as two discrete frequencies.

It is important to note that considering the finite window case in the analysis of the influence of defects on the orientation estimation is a simplified approach. It does not take into consideration the effect of truncation of the Gaussians. For instance, we assumed that the entire oriented defect is contained within the window. In realistic scenarios usually only a segment is fit within the window. To be more precise one needs to represent the signal in spatial domain as a product of \( s'(x,y) \) times a rectangular window. This leads to rather complicated expressions in the Fourier domain and the simplicity of the model is lost. The convolution of the signal with the sinc function in the frequency domain results in limited resolution and needs to be accounted for when one is interested in performance analysis.
CHAPTER 7

CONCLUSIONS AND FUTURE WORK

The impact of low quality on classification should not be ignored. Bayes’ theorem allows incorporating the quality in a form of a conditional distribution of a feature given the raw pattern. This can be formulated as classification in the presence of noisy features [Duda et al., 2001] or as the belief propagation in the case of uncertain evidence [Korb and Nicholson, 2004]. We have proposed to define quality as an uncertainty of the features extracted from the pattern. It is assumed that the uncertainty can be derived from data and based on the knowledge of the feature extraction method. The uncertainty of the quantities estimated from data originates from distortion of the original pattern. These distortions are then propagated through the subsequent processing stages. Ideally, the uncertainty could correspond to a probability function. Examples from the literature suggest that using some form of quality information in order to weight the classifier’s decision yields classification improvement even if the true probability is not known. Such ad hoc approaches likely provide room for improvement. The framework presented in Chapter 2 using a Markov dependence graph can be used in order to propagate the uncertainty to the classification stage. In this work we have shown different ways in which uncertainty can be quantified in the case of the four representative orientation estimation methods. The open question still is how to design the higher level feature extraction mechanisms so that they incorporate the uncertainty of the lower level features. The problem arises from the fact that the higher level features represent different abstraction level. The simple weighting approach of the OFFC curves was tried while working on this thesis but it is yet another ad hoc approach and therefore was not pursued further.

Using datasets from the Fingerprint Verification Competition (FVC), we have shown that smoothing of the orientation field improves the categorization accuracy. It is evident in the case of a low quality dataset. Smoothing has been shown previously to be a useful preprocessing technique that leads to reduction of false singularity detection. Unfortunately, smoothing in our setting also leads to alterations of the orientation field in high curvature regions. It causes loss of fine details and decreases accuracy of the localization of singularities as shown in Chapter 4. We
have proposed an alternative method based on MRF smoothing of the orientation field in which the uncertainty of the orientation estimate is used as a weight in the iterative smoothing. The method is advantageous because of its efficiency. It requires only weighted averaging instead of the optimization method described in [Dass, 2004]. The orientation field estimates obtained this way are satisfactory. It has been confirmed by testing the categorization accuracy using datasets of fingerprint impressions from live sensors. Additional testing using NIST databases could be conducted in the future to compare the performance with the results reported in the literature.

In Chapter 5 we have proposed a set of scalar measures that can be used to quantify the quality. It is accepted to use a coherence measure to determine the reliability of local orientation estimates. This is applicable for the MSE estimate of orientation that can be obtained using PCA of gradients. When the orientation estimate is found from the maximum of the orientation signature, it is more meaningful to use the properties of the entire orientation signature. We have proposed localized scalar metrics that require the orientation signature s(θ) and/ or the probabilistic noise model. These localized metrics can then be then used to compute a global scalar measure of quality for the fingerprint, which is useful for comparison of two impressions. In addition, the spatial location of important points in the fingerprint can be used to properly weight the quality descriptors for better fit in the classification scheme, similarly as in [Tabassi et al., 2004]. The global scalar metrics are computed in a form of quantiles of the distribution of the local descriptors over the entire image.

In Chapter 6 it has been shown that the Fourier Transform representation can be used to illustrate the different approaches to fingerprint ridge line orientation estimation that have been presented in Chapter 5. These different methods essentially present different points of view but the underlying principle used in all of them is common. If one ignores the bandpass filtering aspect, then there are two types of approaches: MMSE and maximum energy. In Chapter 6 we have focused only on the methods that find the orientation estimate as the orientation of maximum energy, since such an approach allows for more detailed analysis of the local orientation through the orientation signature. The MSE methods are suited best in the cases of patterns with only a single orientation. In more complex scenarios involving multiple maxima in an orientation signature or asymmetric distribution of energy, it is more informative to use the orientation signatures and examine the local maxima as possible candidate orientations. Since we are interested in estimation in low quality, it is reasonable to assume that the orientation of maximum energy may be not the true orientation. In such a case the orientation estimation can be
reinforced by using information from the neighborhood. The improvement in orientation estimation in both cases (of MSE and maximum energy methods) can be achieved when spatial frequency-matched filters are used as the preprocessing step. Special care needs to be taken in singular regions and minutiae points. If only the orientation is of interest, then even severe defects, such as missing fragments of ridges, can be “repaired” through matched filtering. This, however, comes at a cost of suppression of fine details such as minutiae. This effect can manifest itself even more severely in the singular regions. This undesired effect of suppression of fine details can be limited in practice by increasing filter’s bandwidth. This phenomenon is explained through the uncertainty principle, i.e. the fact that one can not localize the operator’s energy simultaneously in spatial and spatial frequency domains. The presentation in Chapter 6 has been limited to the methods seeking the maximum energy orientation: Fourier Transform – based and rotating spatial window – based methods. However, it can similarly be conducted for MSE methods.

The defect model presented in Chapter 3 simplifies the analysis of the effect of the quality impairments on the orientation estimation. Based on the qualitative analysis, a defect model has been proposed. In the context of this work, the quality defects can be precisely defined as the root causes of distortions of the otherwise smooth orientation field. In Chapter 6 the most critical ones have been investigated: the additive Gaussian noise and the multiplicative, oriented defects. Other defects are not as crucial to be investigated. For instance, the images with decreased contrast and shifted mean intensity can be normalized and then treated the same way as the examples studied earlier in this chapter. We can expect that as the result of normalization the noise would be scaled up. A special comment is needed to address the case of the model of pressure with which the fingerprint is pressed against the sensor. As it has been shown in Chapter 3, the ideal signal is represented as a sine wave. Such a signal has a localized spectrum. The manifestation of the pressure in the spatial domain, according to the model, is distortion of the sine wave – the ridge width is increased if pressure is increased, and the valley width is increased if pressure is reduced. In the Fourier domain it corresponds to introduction of higher order harmonics. This phenomenon justifies methods that use the entire spectrum (Fourier Transform-based, spatial window PCA of gradients). The model is flexible and yet it is simple enough to allow for analysis.

The future work in the area of the impact of quality on the classification of fingerprint patterns, and in general classification problems should lead toward a formulation of the problem
allowing for propagation of quality as shown in Chapter 2. The uncertainty models for higher level features appear to be crucial and remain to be developed.

The application of the defect model can lead to better and more predictable synthetic pattern generators. This, in turn, could facilitate the design and testing of the fingerprint recognition systems more efficiently. The model presented in Chapter 3 has been developed without considering the physical phenomena. The properties of the skin and the pressure distribution lead to specific characteristics of the appearance of the impressions. Such models are available to a limited extent. The pressure model based on finite element analysis has been presented by [Gerling and Thomas, 2005], and could probably be used in the context of modeling of the pressure across the fingertip that is in contact with the sensing surface. Masks can be used to control the amount of defects such as noise and variations of the skin surface (modeled by Gaussian functions) in order to achieve more realistic effects than those presented in this work. The elastic deformations have not been included in the model. Their effect is especially important in the case of matching and is not as critical in the case of the categorization that was used as the case study throughout this thesis. Such models have been presented by [Cappelli et al., 2001].
REFERENCES


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


[Cappelli et al., 2004b] R. Cappelli, "SFinGe: an Approach to Synthetic Fingerprint Generation", in proceedings of International Workshop on Biometric Technologies (BT2004), Calgary, Canada, pp.147-154, June 2004


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


145


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
APPENDIX

COHERENCE OF GRADIENTS AND COHERENCE OF NORMALIZED GRADIENTS

In [Jiang, 2005], a method for anisotropy estimation was presented involving normalization of gradients. In the course of this work a coherence of normalized gradients was considered as a quality descriptor. We would like to point out two important observations. The first one is related to using multiple descriptors for segmentation or quality assessment. Normalization increases correlation between coherence and standard deviation of intensities, as shown in Figure A.1.

![Figure A.1. Comparison; scatter diagram of standard deviation of intensities and c1 a), and c2 b).](image)

Another observation is that coherence of normalized gradients has almost always a lower value than coherence of gradients, as shown in the figure Figure A.2.
We would like to investigate this further. The following definition of the coherence has been presented in [Bazen and Gerez, 2002].

\[
c_1 = \frac{\sum_{k=1}^{N} g_{sk}}{\sqrt{\sum_{k=1}^{N} |g_s|^2}}
\]

The coherence of normalized gradients can, therefore, be written as

\[
c_2 = \frac{\sum_{k=1}^{N} g_{sn}}{N}
\]

where \( g_{sn} \) is a squared gradient calculated using unit-length gradients.

It is interesting to investigate further the relationship between \( c_1 \), the coherence of non-normalized gradient vectors, and \( c_2 \), the coherence of normalized gradient vectors.

We would like to investigate whether the following relation holds:

\[ c_2 \leq c_1 \]
i.e.

\[
\frac{\sum_{k=1}^{N} g_{sx}^2}{N} \leq \frac{\sum_{k=1}^{N} g_{s}^2}{N} \leq \frac{\sum_{k=1}^{N} |g_{s}|}{N}
\]

The following hold:

1) Gradients calculated on integers numbers yield, \( g_{sx}, g_{sy} \in \{0,1,2\ldots C_{\text{max}}\} \).

2) \( 0 \leq c_1 \leq 1, \ 0 \leq c_2 \leq 1 \)

Local orientation of the fingerprint ridges can be estimated in the mean-squared-error sense using PCA of the gradient field. In this approach, the orientation is defined as an orientation of the eigenvector corresponding to the smallest eigenvalue representing the orientation normal to the orientation of gradient vectors.

In order to apply the PCA method, we assume that \( E[g_{sx}] = E[g_{sy}] = 0 \). The gradient field in a local neighborhood can be characterized by a covariance matrix:

\[
\sum = \begin{bmatrix}
\sigma_{x}^2 & \sigma_{xy}^2 \\
\sigma_{xy}^2 & \sigma_{y}^2
\end{bmatrix}
\]

Let \( \lambda_1 \) and \( \lambda_2 \) be the eigenvalues of the covariance matrix shown above, such that \( \lambda_1 \geq \lambda_2 \) and \( \lambda_1, \lambda_2 \) are real, and positive. The coherence \( c_1 \) can then be defined:

\[
c_1 = \frac{\lambda_1^{(1)} - \lambda_2^{(1)}}{\lambda_1^{(1)} + \lambda_2^{(1)}}
\]

The eigenvalues are equal to the variances along the principal axes. We expect that after normalization of the gradient vectors the variances will decrease. It is the consequence of the assumption that the x and y components of gradients can assume integer values from 0 to 255. The smallest nonzero gradient therefore has magnitude equal to 1, in which case normalization will not have any effect.
For normalized gradients the PCA looks as follows. As previously done, we obtain 2 principal components, and 2 eigenvalues: $\lambda_1 \geq \lambda_2$.

The coherence of normalized gradient field is

$$c_2 = \frac{\lambda^{(2)}_1 - \lambda^{(2)}_2}{\lambda^{(2)}_1 + \lambda^{(2)}_2}$$

Since normalized gradient vectors have x and y components that are less than or equal to gradient vector components prior to normalization, we can write

$$\lambda^{(1)}_1 = k_1 \lambda^{(2)}_1$$

$$\lambda^{(1)}_2 = k_2 \lambda^{(2)}_2$$

where $k_i \in [0,1]$, $i=1,2$

We can now write $c_1$ as a function of the principal values of normalized gradient vector field:

$$c_1 = \frac{\lambda^{(1)}_1 - \lambda^{(1)}_2}{\lambda^{(1)}_1 + \lambda^{(1)}_2} = \frac{k_1 \lambda^{(2)}_1 - k_2 \lambda^{(2)}_2}{k_1 \lambda^{(2)}_1 + k_2 \lambda^{(2)}_2} = \frac{\lambda^{(2)}_1 - k_2 \lambda^{(2)}_2}{k_1 \lambda^{(2)}_1 + k_2 \lambda^{(2)}_2}$$

From the formula above it can be seen that $c_1 \geq c_2$, given $k_1 \geq k_2$. Experimental data show that this need not always hold, but it seems to be the case most of the time. In rare cases $c_2 \geq c_1$, which is possible since the normalization can lead to rotation of the principal axes. In other words, it can occur when the orientations of the principal vectors of the distribution of gradients before and after normalization are different.

$$\angle e^{(1)}_1 \neq \angle e^{(2)}_1$$
\[ \angle e_2^{(1)} \neq \angle e_2^{(2)} \]

The comparison of the two measures based on experiments show that \( c_1 \geq c_2 \) may be related to the fact that normalization of gradients amplifies the effect of weak gradients (\( c_2 \) is more sensitive to inconsistencies of local gradient orientations). The weak gradients are mostly attributed to noise and have random orientations which may explain the relation \( c_1 > c_2 \).