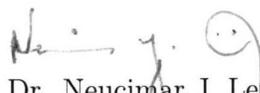


Transformações Multi-escala para Segmentação de Imagens de Impressões Digitais

Raoni Florentino da Silva Teixeira

Este exemplar corresponde à redação final da
Dissertação devidamente corrigida e defendida
por Raoni Florentino da Silva Teixeira e apro-
vada pela Banca Examinadora.

Campinas, 24 de março de 2011.


Prof. Dr. Neucimar J. Leite
Instituto de Computação - UNICAMP
(Orientador)

Dissertação apresentada ao Instituto de Com-
putação, UNICAMP, como requisito parcial para
a obtenção do título de Mestre em Ciência da
Computação.

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO IMECC DA UNICAMP**

Bibliotecária: Maria Fabiana Bezerra Müller – CRB8 / 6162

Teixeira, Raoni Florentino da Silva

T235t Transformações multi-escala para segmentação de imagens de impressões digitais/Raoni Florentino da Silva Teixeira-- Campinas, [S.P. : s.n.], 2011.

Orientador : Neucimar Jerônimo Leite.

Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Computação.

1.Biometria. 2.Processamento de imagens. 3.Morfologia matemática. 4.Simplificação de imagem. I. Leite, Neucimar Jerônimo. II. Universidade Estadual de Campinas. Instituto de Computação. III. Título.

Título em inglês: Multi-scale transformations for fingerprint image segmentation

Palavras-chave em inglês (Keywords): 1.Biometrics. 2.Image processing. 3.Mathematical morphology. 4.Image simplification.

Área de concentração: Processamento de imagens

Titulação: Mestre em Ciência da Computação

Banca examinadora: Prof. Dr. Neucimar Jerônimo Leite (IC – UNICAMP)
Prof. Dr. Hélio Pedrini (IC – UNICAMP)
Prof. Dr. Aparecido Nilceu Marana (Fac.de Ciências de Bauru-UNESP)

Data da defesa: 24/03/2011

Programa de Pós-Graduação: Mestrado em Ciência da Computação

TERMO DE APROVAÇÃO

Dissertação Defendida e Aprovada em 24 de março de 2011, pela Banca examinadora composta pelos Professores Doutores:



Prof. Dr. Aparecido Nilceu Marana
Faculdade de Ciências de Bauru / UNESP



Prof. Dr. Hélio Pedrini
IC / UNICAMP



Prof. Dr. Neucimar Jerônimo Leite
IC / UNICAMP

Transformações Multi-escala para Segmentação de Imagens de Impressões Digitais

Raoni Florentino da Silva Teixeira¹

Março de 2011

Banca Examinadora:

- Prof. Dr. Neucimar J. Leite
Instituto de Computação - UNICAMP (Orientador)
- Prof. Dr. Hélio Pedrini
Instituto de Computação - UNICAMP
- Prof. Dr. Aparecido Nilceu Marana
Faculdade de Ciências de Bauru - UNESP
- Prof. Dr. Anderson de Rezende Rocha (Suplente)
Instituto de Computação - UNICAMP
- Prof. Dr. Arnaldo de Albuquerque Araújo (Suplente)
Departamento de Ciência da Computação - UFMG

¹Suporte financeiro de: CNPq, 03/2009 à 02/2011

Resumo

A identificação baseada em impressões digitais tem recebido considerável atenção nos últimos anos devido à crescente procura pela identificação automática de indivíduos, tanto em aplicações forenses quanto empresariais, por exemplo. Uma importante etapa que deve ser considerada nessas aplicações é a *segmentação* da imagem que constitui a impressão digital. Nesse contexto, o termo *segmentação* refere-se à separação da imagem em duas regiões, denominadas área da impressão (*foreground*) e fundo (*background*), a fim de evitar que características utilizadas no reconhecimento e/ou classificação das impressões digitais correspondentes sejam extraídas de regiões impróprias. Normalmente, as abordagens de segmentação encontradas na literatura não consideram imagens provenientes de diferentes bases de dados (ou sensores), em virtude da diversidade das propriedades e características encontradas em cada sensor e, em geral, o desempenho dos métodos existentes é baixo quando lidam com bases de dados heterogêneas. Neste sentido, a segmentação de imagens oriundas de diferentes sensores constitui um problema ainda a ser explorado. Este trabalho apresenta um conjunto de transformações de imagens que pode ser utilizado para esse fim, ou seja, para segmentação de imagens de impressões digitais provenientes de diferentes sensores sem que seja necessário, por exemplo, uma pré-classificação ou treinamento. De modo geral, estas transformações são baseadas em operadores morfológicos do tipo *toggle* que apresentam características interessantes de simplificação de imagens. Os resultados obtidos considerando imagens de diferentes bases de dados mostram que o método proposto supera abordagens bem conhecidas da literatura que representam o *estado-da-arte*.

Abstract

Fingerprint identification has received considerable attention in the last few years, due to an increasing demand for human automatic identification in areas concerning, for example, forensic and business applications. An important step to be considered in such applications is the fingerprint image segmentation. In this context, the term refers to splitting the image into two regions, namely, foreground and background, in order to avoid the extraction of features used in automatic classification and recognition from noisy regions. Usually, the segmentation methods found in the literature do not consider images from different databases (or sensors) and, in a general way, dealing with heterogeneous databases constitutes an open problem not well explored in the literature. This work presents a new set of image transformations related to fingerprint segmentation of images acquired from different sensors without any requirement for pre-classification or training. As we will elsewhere, these transformations are based on morphological toggle operators which present interesting image simplification properties. We evaluate our approach on images of different databases, and show its improvements when compared against other well-known state-of-the-art segmentation methods discussed in literature.

Agradecimentos

Durante o desenvolvimento desse trabalho muitas coisas aconteceram e muitas pessoas me ajudaram. Nesta seção, tento lembrar das pessoas que somaram (algumas até multiplicaram) e agradeço:

- À Deus, por tudo que representa;
- À minha namorada, Ana, pela paciência, amor e suporte sem os quais seria muito difícil passar por este período;
- Aos meus pais, Osvaldo e Tamara, pelo que sou e pelo esforço, amor, carinho, e tudo mais que dispensaram. Em especial ao meu pai, cujo cuidado o fez abandonar suas atividades quando tive de operar.
- Aos meus irmãos, Rauan e Raiani, e aos demais familiares pelo apoio e amor;
- Ao meu orientador, prof. Dr. Neucimar J. Leite, pela orientação na realização deste trabalho e pelas conversas;
- Aos meus antigos orientadores da época da graduação;
- Aos membros da banca examinadora pelas sugestões;
- Aos meus amigos e colegas de república, de laboratório e de grupo de pesquisa, pois sem eles esse período teria sido menos engraçado e feliz;
- Ao pessoal do IC (colegas, professores e demais funcionários) pelo ambiente agradável dos últimos dois anos;
- A todas as outras pessoas que injustamente não foram lembradas, o meu muito obrigado: *“thanks for all the fish!!!”*

Sumário

Resumo	v
Abstract	vi
Agradecimentos	vii
1 Introdução	1
1.1 Principais contribuições	2
1.2 Organização da dissertação	3
2 Trabalhos Correlatos	4
3 Fundamentos Teóricos	7
3.1 Características de uma impressão digital	7
3.2 Morfologia Matemática	10
3.2.1 Algumas propriedades	10
3.2.2 Elementos e funções estruturantes	12
3.2.3 Operações básicas	13
3.2.4 Outras operações	15
3.3 Operadores direcionais	17
3.3.1 Método do gradiente	18
3.3.2 Operador multi-escala direcional	19
4 Método Proposto	22
4.1 Descrição geral do método	22
4.2 Aspectos relacionados ao contraste	26
4.3 Parâmetros <i>versus</i> estruturas da imagem	29
4.4 O operador toggle e a teoria espaço-escala	31
4.5 Extensões	33

5 Experimentos	35
5.1 Medidas de avaliação	35
5.2 Experimento 1	36
5.3 Experimento 2	39
5.4 Experimento 3	42
5.5 Experimento 4	43
5.6 Experimento 5	43
6 Conclusão e Trabalhos Futuros	47

Lista de Tabelas

2.1	Comparação entre métodos de segmentação de impressões digitais. NS, S, P e B significam <i>não supervisionado</i> , <i>supervisionado</i> , <i>segmentação pontual</i> e <i>por bloco</i> , respectivamente.	5
5.1	Características das imagens pertencentes a base de dados FVC	37
5.2	Médias da função F-score obtidas para bases de dados FVC	38
5.3	Médias da função de <i>Jaccard</i> obtidas para bases de dados FVC	38
5.4	Médias da função F-score obtidas para a base de dados do NIST.	41
5.5	Médias da função de <i>Jaccard</i> obtidas para a base de dados do NIST.	42
5.6	Médias da função F-score obtidas para base de dados FVC2004 - DB3.	42
5.7	Média da função <i>Jaccard</i> obtida para base de dados FVC2004 - DB3.	43
5.8	Médias da função F-score obtidas para bases de dados FVC2004.	43
5.9	Médias da função de <i>Jaccard</i> obtidas para bases de dados FVC2004.	45
5.10	Médias da função F-score obtidas para bases de dados FVC.	45
5.11	Médias da função de <i>Jaccard</i> obtidas para bases de dados FVC.	46

Lista de Figuras

3.1	Cristas e vales em uma impressão digital [Maltoni et al., 2009].	8
3.2	Núcleo e delta em uma imagem de impressão digital.	8
3.3	Da esquerda para a direita: arco plano, arco angular, laço à esquerda, laço à direita, misto e verticilo.	9
3.4	Imagem de impressão digital com três minúcias. As cristas em detalhe em (1) e (2) correspondem às características conhecidas como terminações e a crista apresentada em (3) corresponde a uma bifurcação.	9
3.5	Imagem de impressão digital apresentando poros.	10
3.6	Exemplos de funções estruturantes planares contidos em uma janela 3×3 . O ponto branco denota o centro do elemento estruturante.	12
3.7	Exemplos de erosão e dilatação: (a) imagem original, (b) erosão considerando uma função estruturante planar 10×10 e (d) dilatação considerando esta mesma função.	14
3.8	Exemplos de abertura e fechamento: (a) imagem original, (b) resultado da abertura utilizando um elemento estruturante circular de raio $r = 20$ e (c) fechamento utilizando um elemento estruturante circular de raio $r = 5$	16
3.9	Exemplos de filtragem morfológica de uma imagem com ruído do tipo sal-e-pimenta: (a) imagem original com ruído, (b) resultado da abertura utilizando uma função estruturante circular de raio $r = 3$, (c) imagem obtida após o fechamento, utilizando essa mesma função estruturante e (d) resultado da aplicação de um fechamento seguido de uma abertura.	16
3.10	Detalhe do contraste obtido entre a informação na direção das papilas e sua direção perpendicular.	19
3.11	Exemplo de pontos contidos em segmento de linha de direção igual a 45° e tamanho 9.	20
4.1	Fluxograma do método proposto.	23
4.2	Exemplo de segmentação para uma imagem do FVC2000: (a) imagem original e (b) resultado da aplicação da equação 4.1.	24
4.3	Estimativa de background computada utilizando $k = 15$ e $\sigma = 20$	24

4.4	Exemplo dos passos realizados na segmentação de uma imagem do FVC2004: (a) imagem original, (b) orientação fina do campo direcional, (c) estimativa do background (d) resultado do processo de subtração do fundo, (e) orientação grosseira do campo direcional computada considerando a imagem de subtração do fundo (nesta figura, cada um dos tons de cinza indica uma direção diferente) e (f) imagem final segmentada.	25
4.5	Relação de proximidade entre os valores transformados e a imagem original: (a) imagem original, (b) mapa de proximidade considerando a escala $\sigma = 10$ e número de iterações $k = 10$, (c) mapa de proximidade considerando o mesmo número de iterações e a escala $\sigma = 20$ e (d) mapa de proximidade considerando a escala $\sigma = 30$ e o mesmo número de iterações.	26
4.6	Representação 3D de uma impressão digital ideal.	27
4.7	Relação de proximidade para uma imagem de impressão digital: (a) imagem original, mapa de proximidade considerando o número de iterações $k = 15$ e escalas (b) $\sigma = 5$, (c) $\sigma = 10$, (d) $\sigma = 20$, (e) $\sigma = 30$, (f) $\sigma = 40$ e (g) $\sigma = 50$	28
4.8	Mapa de proximidade utilizando elementos estruturantes planares: (a) imagem original, (b) mapa de proximidade com $k = 15$, (c) imagem original e (d) mapa de proximidade com o mesmo número de iterações k	29
4.9	Exemplo de segmentação para uma imagem pertencente à base de dados do NIST com uma região de borramento: (a) corresponde à estimativa do fundo (Equação 4.2) e (b) ao resultado da segmentação.	30
4.10	Exemplo de aplicação do mapeamento da Equação 4.2 para uma mesma escala $\sigma = 20$ e diferentes números de iterações: (a) imagem original, (b) resultado utilizando número de iterações $k = 1$, (c) $k = 5$, (d) $k = 10$ e (e) $k = 15$	31
4.11	Função estruturante pirâmide.	32
4.12	Fluxograma do (a) método de Bazen e Gerez [2001] original e da (b) extensão proposta.	34
5.1	Exemplo de imagens retiradas das bases de dados do FVC2004.	37
5.2	Comparação entre métodos: (a) e (c) segmentação obtida pelo método proposto e (b) e (d) segmentação pelo método de Bazen e Gerez [2001].	39
5.3	Exemplo de segmentação obtida pelo método proposto para imagens das bases de dados do FVC.	40
5.4	Exemplos de imagens da base de dados do NIST.	41

5.5	Exemplos de segmentação obtida, respectivamente da direita para esquerda, com o sistema biométrico do NIST, o algoritmo de Bazen e Gerez [2001] e o método proposto.	41
5.6	Exemplos de segmentação com borramento: (a) método proposto e (b) método do NIST.	42
5.7	Exemplos de segmentação pela modificação proposta: (a), (d), (g) e (j) imagens originais, (b), (e), (h) e (k) modificação proposta e (c), (f), (i) e (l) método de Bazen e Gerez [2001].	44

Capítulo 1

Introdução

Reconhecimento biométrico, ou simplesmente biometria, tem recebido considerável atenção nos últimos anos devido à crescente procura pela identificação automática de indivíduos, tanto em aplicações forenses quanto empresariais, por exemplo [Maltoni et al., 2009].

O foco deste trabalho é impressões digitais, que de acordo com a literatura, é a característica biométrica mais estudada e difundida, utilizada em larga escala em investigações criminais e aplicações comerciais que exigem controle de acessos e identificação *online*. Este tipo de característica tem sido considerado popularmente como sinônimo de biometria [Maltoni et al., 2009].

A maioria destas aplicações possui uma etapa de pré-processamento em que ruídos das imagens são filtrados e a própria impressão é segmentada com relação ao fundo ou *background*. Tal etapa é importante para evitar que características (*pontos singulares* ou *minúcias*, dependendo do nível de avaliação) sejam extraídas de regiões ruidosas ou do *background*, prejudicando assim sua análise e reconhecimento.

No entanto, como a impressão digital pode ser obtida de diversas formas, através, por exemplo, de um sensor capacitivo, óptico, térmico ou de campo elétrico, ou ainda, simplesmente utilizando tinta e papel, e as imagens obtidas por cada um deles possuem propriedades diferentes, é difícil definir um método de pré-processamento e segmentação robusto e genérico.

Nesse sentido, os melhores resultados de filtragem e segmentação conhecidos na literatura são definidos a partir de abordagens supervisionadas que requerem uma fase inicial de aprendizado por classe [Bazen e Gerez, 2001, Maltoni et al., 2009], em que várias imagens obtidas com um mesmo sensor, e portanto da mesma classe, devem ser segmentadas manualmente para fins de treinamento. Esta dependência de classes é ruim, já que obriga uma pré-classificação manual de todas as imagens analisadas.

Tal dependência é descrita na literatura a partir do conceito de *interoperabilidade de*

sensor que se refere à capacidade de um sistema biométrico se adaptar ou não a dados (ou imagens, no caso de impressões digitais) provenientes de diferentes sensores ou bases de dados [Ross e Jain, 2004].

Um exemplo típico em que a interoperabilidade é interessante e em certo ponto até crucial consiste em uma aplicação que considera impressões digitais obtidas através de sensores modernos e utilizando tinta e papel. Historicamente, as impressões digitais eram (e, ainda, são em algumas ocasiões) colhidas utilizando tinta e papel, porém, avanços na tecnologia dos sensores popularizaram e difundiram sistemas biométricos baseados nos mesmos. Isto, dentre outros fatores, contribuiu para a criação de sistemas biométricos interoperáveis que se relacionam com imagens obtidas por ambas as técnicas.

Em geral, como a maioria dos métodos utilizados por sistemas biométricos são desenvolvidos considerando somente um tipo de sensor ou base de dados, eles sofrem portanto de problemas relacionados à *interoperabilidade*, de tal forma que o desempenho desses sistemas é baixo quando lidam com bases de dados heterogêneas [Guo et al., 2009, Ross e Jain, 2004].

Alguns trabalhos recentes apresentam avanços nesse sentido [Ross e Nadgir, 2006, Maltoni et al., 2009]. No entanto, sobretudo no que se refere à segmentação de imagens de impressão digital, o assunto é pouco abordado na literatura [Yang et al., 2010, Guo et al., 2009].

Este trabalho tem como objetivo principal, a definição de um conjunto de transformações de imagens que possam ser utilizadas no pré-processamento e segmentação de imagens de impressões digitais sem que seja necessário uma pré-classificação ou treinamento, o que significa uma independência com relação aos sensores utilizados.

Estas transformações são realizadas tendo por base operações de morfologia matemática e a informação direcional obtida da imagem de impressão digital. Mais especificamente, foram exploradas as propriedades de simplificação de um operador de mapeamento do tipo *toggle* e do operador multi-escala direcional [Oliveira e Leite, 2008] na definição de uma transformação de filtragem.

1.1 Principais contribuições

As principais contribuições desse trabalho são as seguintes:

- definição e avaliação de um operador morfológico do tipo *toggle* com propriedades de filtragem significativas para segmentação de imagens de impressões digitais;
- estudo comparativo de algumas técnicas de segmentação de imagens de impressões digitais considerando diferentes bases de dados;

- definição, aplicação e teste de um método para segmentação de imagens de impressões digitais independente de classes de sensores (não requer treinamentos específicos) e
- segmentação e anotação manual de um conjunto de 878 imagens de diferentes sensores, utilizado como *ground truth*.

1.2 Organização da dissertação

O texto desta dissertação de mestrado está organizada em seis capítulos, da seguinte maneira. O Capítulo 2 apresenta uma revisão bibliográfica, contendo uma breve descrição dos principais métodos de segmentação de impressões digitais encontrados na literatura. No Capítulo 3 são apresentados os fundamentos teóricos utilizados neste trabalho. O Capítulo 4 descreve o método de segmentação proposto, discutindo seus passos e extensões. Em seguida, o Capítulo 5 mostra os resultados dos testes experimentais realizados e o Capítulo 6 apresenta as conclusões e os trabalhos futuros.

Capítulo 2

Trabalhos Correlatos

O problema geral abordado nesta dissertação é a segmentação de imagens de impressão digital. Mais especificamente, este trabalho trata da segmentação de imagens obtidas através de diferentes sensores, o que é conhecido na literatura como *interoperabilidade de sensor*.

Este capítulo apresenta uma breve revisão do *estado-da-arte*, discutindo algumas das principais abordagens encontradas na literatura. Informações adicionais podem ser encontradas em [Maltoni et al., 2009], que apresenta uma revisão cronológica sobre o tema.

Em virtude de sua importância, várias abordagens para segmentação de imagens de impressões digitais podem ser encontradas na literatura, como ilustra a Tabela 2.1. Em geral, tais abordagens podem ser globalmente classificadas em *supervisionadas* ou *não-supervisionadas*, dependendo de como o treinamento ou a definição de parâmetros é realizada, e em *segmentação pontual* ou *por blocos*, dependendo se são classificados *pixels* ou blocos (*conjuntos de pixels*) das imagens.

Pode-se ainda adotar uma classificação baseada na capacidade do método de se adaptar ou não a dados obtidos de vários sensores. Como mencionado anteriormente, tal característica é conhecida na literatura como *interoperabilidade de sensor* e tem recebido pouca atenção, pois sistemas biométricos consideram geralmente *dados* provenientes de um tipo de sensor. Isso, porém, dificulta a interação entre sensores (ou base de dados) e limita a usabilidade do sistema [Ross e Jain, 2004].

Mehre et al. [1987] propõem um método de segmentação por bloco de acordo com o histograma das orientações das papilas. A imagem de orientações é dividida em blocos de 16×16 *pixels* e o histograma é computado para cada bloco. A presença de um pico significativo neste histograma denota um padrão orientado e histogramas uniformes caracterizam regiões do fundo da imagem.

Ratha et al. [1995] descrevem um método de segmentação por bloco baseado na variância obtida considerando um segmento de linha localizado na direção ortogonal ao

Artigos	Interoperabilidade	Aprendizado	Classificação	Abordagens
Mehltre et al. [1987]		NS	B	análise do histograma das orientações
Ratha et al. [1995]		NS	B	Variância na direção perpendicular à orientação
Shen et al. [2001]		NS	B	Convolução com um conjunto de filtros de Gabor
Bazen e Gerez [2001]		S	P	Rede neural considerando coerência, média e variância.
Chen et al. [2004]		S	B	Rede Neural considerando grau de agrupamento, média e variância.
Zhu et al. [2006]		S	B	Rede neural em uma espaço com onze características.
Watson et al. [2007]		NS	B	Binarização e remoção de componentes
Yang et al. [2010]	✓	NS	B	K-means considerando coerência, média e variância.

Tabela 2.1: Comparação entre métodos de segmentação de impressões digitais. NS, S, P e B significam *não supervisionado*, *supervisionado*, *segmentação pontual* e *por bloco*, respectivamente.

campo direcional. Tal informação gera um mapa com contraste direcional da impressão digital que pode ser usado também como descritor de qualidade, no qual regiões de baixa qualidade (borradas e de fundo, por exemplo) possuem valores pequenos e regiões com informação direcional bem definida possuem valores elevados.

Shen et al. [2001] consideram um conjunto de filtros de Gabor, representando diferentes direções, e usam o desvio padrão obtido entre as informações destas direções como critério para segmentação. Tal abordagem baseia-se na observação de que, ao contrário do fundo da imagem, na região da impressão digital há uma direção dominante e, portanto, o desvio padrão nesta região deve ser alto.

Bazen e Gerez [2001] propuseram um método de segmentação pontual, que utiliza três descritores (média, variância e coerência) e um classificador linear para rotular os pixels correspondentes ao fundo e à área da impressão digital. Uma abordagem supervisionada é utilizada para treinar um classificador baseado em rede neural e um pós-processamento morfológico (abertura seguida de fechamento) é realizado para eliminar buracos e regu-

larizar a silhueta da área da impressão digital. Devido à sua especificidade, no que diz respeito a base de dados considerada (ou sensor), esta abordagem produz resultados muito precisos de segmentação e é a base de várias técnicas discutidas na literatura, envolvendo filtragem e análise [Yang et al., 2010, Chen et al., 2004, Maltoni et al., 2009].

Chen et al. [2004] propuseram um método não supervisionado de segmentação por bloco. Um classificador linear é treinado considerando três características (grau de agrupamento, média e variância) e um pós-processamento morfológico (abertura seguida de fechamento) é realizado para reduzir o número de erros ocorridos nessa classificação.

Zhu et al. [2006] apresentaram um método de segmentação por bloco baseado na avaliação da correção do campo direcional, sendo os blocos com orientação *incorreta* classificados como *background*. Uma rede neural é treinada considerando 11 descritores, incluindo a magnitude do gradiente, a variância e a distância entre as cristas da impressão digital.

Watson et al. [2007] descrevem o *sistema de identificação de impressão digital* desenvolvido pelo NIST (*National Institute of Standards and Technology*). O método de segmentação contido nesse sistema primeiro binariza a imagem e então realiza um pós-processamento morfológico para preservar apenas o maior componente conexo da imagem binarizada. Apesar da simplicidade dessa cadeia de processamento, a qualidade dos resultados obtidos nas impressões digitais roladas do NIST é alta (*F-measure* em torno de 0.95).

Recentemente, Yang et al. [2010] apresentaram um método de segmentação que considera o problema da interoperabilidade de sensor. O método define um agrupamento de blocos baseado no classificador k-means a partir de um vetor de características composto pelas informações estatísticas de média, variância e coerência. Além disso, um pós-processamento morfológico (abertura seguida de fechamento) é realizado para melhorar o resultado da classificação.

Capítulo 3

Fundamentos Teóricos

Este capítulo apresenta os conceitos básicos relacionados às áreas de biometria e processamento de imagens considerados neste trabalho. O mesmo é organizado da seguinte maneira: a Seção 3.1 apresenta uma pequena introdução à biometria e as principais características das imagens de impressão digital consideradas neste trabalho. Na Seção 3.2 são apresentados alguns conceitos básicos sobre morfologia matemática, incluindo informações sobre transformações dependentes de escala. Por fim, a Seção 3.3 apresenta o problema de estimativa do *campo direcional* e alguns operadores conhecidos na literatura.

3.1 Características de uma impressão digital

Algumas das principais características consideradas no reconhecimento biométrico são:

- características *fisiológicas* e *anatômicas*, que correspondem aos atributos inatos aos indivíduos como, por exemplo, DNA, formato da orelha, face, termograma facial, impressão digital, geometria da mão, íris, veias da mão, odor, retina e voz;
- características *comportamentais*, que correspondem a ações e atos dos indivíduos, tais como caminhar e assinatura e que podem sofrer alterações em virtude de doenças ou idade.

Em geral, as características fisiológicas são mais confiáveis e precisas do que as comportamentais, pois são mais difíceis de serem copiadas ou imitadas, além de sofrerem menos alterações com o passar do tempo. Este trabalho considera apenas as características fisiológicas, concentrando-se particularmente nas impressões digitais que constituem as características mais difundidas e estudadas atualmente [Maltoni et al., 2009].

Uma impressão digital é composta por cristas e vales, conforme ilustra a Figura 3.1, e sua definição é o resultado do contato da superfície dos dedos das mãos com um deter-

minado objeto (quando uma pessoa toca algo com os dedos, geralmente deixa resíduos visíveis e invisíveis na superfície do que foi tocado).



Figura 3.1: Cristas e vales em uma impressão digital [Maltoni et al., 2009].

O resíduo é uma cópia da impressão digital e pode ser coletado para posterior estudo e comparação de duas formas diferentes. Para tanto, pode-se considerar, por exemplo, um sensor eletrônico ou simplesmente tinta e papel.

Impressões digitais foram classificadas de diversas maneiras, ao longo da história, e de acordo com a literatura são analisadas atualmente sob níveis de singularidade globais, locais e muito finos [Maltoni et al., 2009, Jain et al., 2007].

No nível global, pode-se observar pontos de singularidade chamados de núcleos e deltas, como mostra a Figura 3.2. O *delta* corresponde ao centro de uma região triangular onde há três diferentes direções dominantes e o *núcleo* ao centro de uma região de alta curvatura. Nesse nível, as impressões digitais são geralmente classificadas em três grandes classes: arco, laço e verticilo [Jain et al., 2007]. Essas classes podem sofrer outras subdivisões: arco plano, arco angular, laço à esquerda, laço à direita, verticilo e mistos, como ilustra a Figura 3.3.

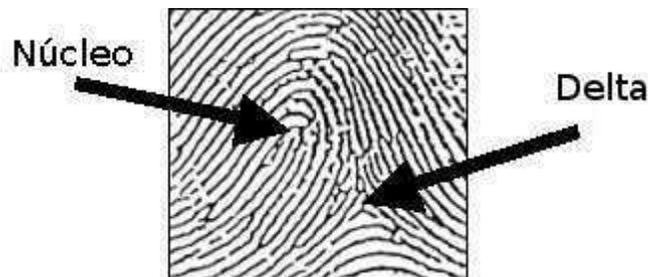


Figura 3.2: Núcleo e delta em uma imagem de impressão digital.

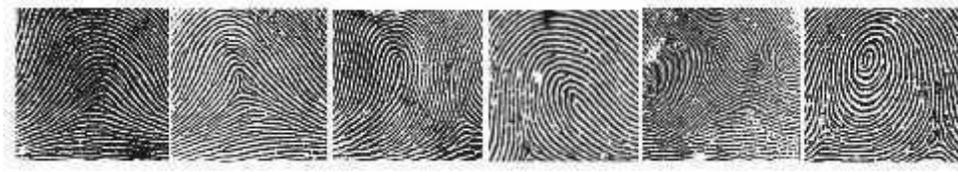


Figura 3.3: Da esquerda para a direita: arco plano, arco angular, laço à esquerda, laço à direita, misto e verticilo.

No nível local, pode-se observar outros detalhes denominados *minúcias*. Os dois principais tipos de minúcias são as bifurcações e terminações apresentados na Figura 3.4. O NIST (*National Institute of Standards and Technology*) propôs uma taxonomia que inclui, além desses dois tipos de minúcias, uma classe denominada *composta* que compreende trifurcações e cruzamentos, e uma outra para minúcias *indeterminadas* [McCabe et al., 2007].

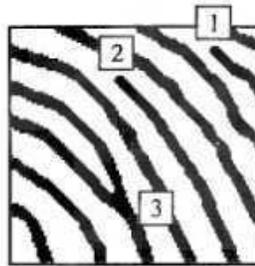


Figura 3.4: Imagem de impressão digital com três minúcias. As cristas em detalhe em (1) e (2) correspondem às características conhecidas como terminações e a crista apresentada em (3) corresponde a uma bifurcação.

Finalmente, no nível muito fino, detalhes tais como os poros dos dedos contidos nas cristas da impressão digital podem ser detectados, como ilustra a Figura 3.5 que mostra uma imagem proveniente da base de dados PolyU HRF ¹. Essa informação pode ser útil ao reconhecimento porém, na prática, as imagens precisam ser de alta resolução, superior a 1000 dpi ² [Jain et al., 2007, 2006, Zhao et al., 2010].

¹Disponível em: <http://www.comp.polyu.edu.hk/~biometrics/HRF/HRF.htm>

²Sigla para *dots per inch*, ou em português, pontos por polegada.

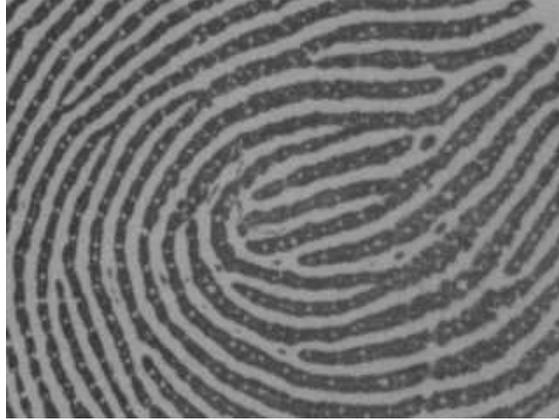


Figura 3.5: Imagem de impressão digital apresentando poros.

3.2 Morfologia Matemática

Morfologia Matemática consiste em uma abordagem não-linear para o processamento de imagens definida a partir de um conjunto de transformações que auxiliam na análise, segmentação e reconhecimento, tendo como base a informação da geometria desta imagem e da forma dos seus objetos.

O resultado de tais transformações depende essencialmente da comparação do conteúdo de uma imagem de interesse com outra menor e de forma conhecida, denominada *elemento* ou *função estruturante*. De maneira geral, esta outra imagem contém características geométricas e/ou topológicas relacionadas com a informação que se pretende extrair da imagem de interesse [Soille, 1999].

Esta seção descreve alguns conceitos relacionados com essa teoria, da seguinte maneira. Inicialmente, na Subseção 3.2.1 são apresentadas algumas das principais propriedades algébricas e topológicas das transformações morfológicas. Na sequência, a Subseção 3.2.2 apresenta considerações sobre diferentes *funções estruturantes*, enquanto a Subseção 3.2.3 descreve as duas operações morfológicas básicas, denominadas *erosão* e *dilatação*, e algumas de suas propriedades. Por fim, a Subseção 3.2.4 apresenta outras transformações importantes, incluindo os operadores do tipo *toggle*.

3.2.1 Algumas propriedades

As transformações morfológicas podem ser consideradas sobre imagens binárias e em tons de cinza, vistas como conjuntos ou funções, respectivamente. A estas imagens pode-se associar operadores booleanos de *união*, \cup , e *intersecção*, \cap , no caso de conjuntos, e os operadores *max* e *min*, no caso de funções.

Tais transformações dispõem de algumas propriedades importantes que permitem definir aspectos de seu comportamento e ajudam na escolha de uma transformação apropriada para um determinado problema [Soille, 1999]. Algumas dessas propriedades são apresentadas a seguir.

Dada uma função ou imagem f , uma transformação morfológica ψ é dita *extensiva* se, e somente se

$$\psi(f) \geq f, \quad (3.1)$$

e *anti-extensiva* se, e somente se

$$\psi(f) \leq f. \quad (3.2)$$

A mesma transformação é dita *crescente* se ela preserva a relação de ordem entre imagens, ou seja,

$$f \leq g \Rightarrow \psi(f) \leq \psi(g). \quad (3.3)$$

Em caso contrário ($f \leq g \Rightarrow \psi(f) \geq \psi(g)$), ela é dita *decrecente*. A transformação ψ é *idempotente* se

$$\psi^{(n)}(f) = \psi(f), \quad (3.4)$$

ou seja, se o resultado de várias iterações do operador ψ sobre a imagem equivale ao resultado de uma única aplicação. Algumas transformações morfológicas utilizam essa propriedade como critério de parada e outras dependem dela para evitar problemas, tais como oscilações.

Duas transformações ψ_1 e ψ_2 são *duais* se, e somente se

$$\psi_1(f) = \psi_2^*(f) \Rightarrow \psi_1(f) = (\psi_2(f^c))^c, \quad (3.5)$$

em que c denota o operador complemento e $*$ o negativo. Em outras palavras, duas transformações ψ_1 e ψ_2 são ditas duais se a aplicação de ψ_1 em uma imagem f é equivalente a considerar o complemento da aplicação de ψ_2 no complemento da imagem f . Caso se configure a *dualidade* entre duas transformações ψ_1 e ψ_2 , algumas das propriedades descritas anteriormente são herdadas diretamente por uma transformação ψ_2 a partir de sua transformação dual ψ_1 e assim

$$\begin{aligned} \psi_1 \text{ idempotente} &\Rightarrow \psi_2 \text{ idempotente}, \\ \psi_1 \text{ crescente} &\Rightarrow \psi_2 \text{ crescente}, \\ \psi_1 \text{ anti-extensiva} &\Rightarrow \psi_2 \text{ extensiva} \\ \psi_1 \text{ extensiva} &\Rightarrow \psi_2 \text{ anti-extensiva} \end{aligned}$$

Finalmente, uma transformação é dita *homotópica* se ela não modifica o número de conexidade N_n , isto é

$$N_n(\psi(f)) = N_n(f). \quad (3.6)$$

Vale lembrar ainda que, na prática, as imagens são processadas no plano discreto \mathbb{Z}^2 , com funções assumindo valores em \mathbb{Z} .

3.2.2 Elementos e funções estruturantes

Para se extrair informações relevantes de uma imagem é fundamental determinar adequadamente a vizinhança a ser considerada no cálculo do valor transformado em cada pixel. Em se tratando de morfologia matemática, isso depende da escolha apropriada de um elemento ou função estruturante, sendo que *elementos estruturantes* estão associados ao caso binário e *funções estruturantes* às imagens em níveis de cinza. De maneira geral, essa função define, entre outros, o tamanho e a forma da região a serem considerados pela transformação morfológica.

Tais funções podem ser classificadas em duas categorias: planares, em que todas as posições possuem valor zero, ou não-planares, quando os valores distintos estão associados a altura da função estruturante. Esses valores distintos podem ser interpretados como pesos atribuídos a diferentes posições da função.

A Figura 3.6 apresenta alguns exemplos de funções estruturantes planares contidos em uma janela 3×3 . A escolha da mais apropriada depende da estrutura da informação que se deseja extrair. Neste trabalho, por exemplo, segmentos de linha direcionais, como o apresentado na Figura 3.6(b), são considerados na extração de algumas informações estruturais que compõem a impressão digital.

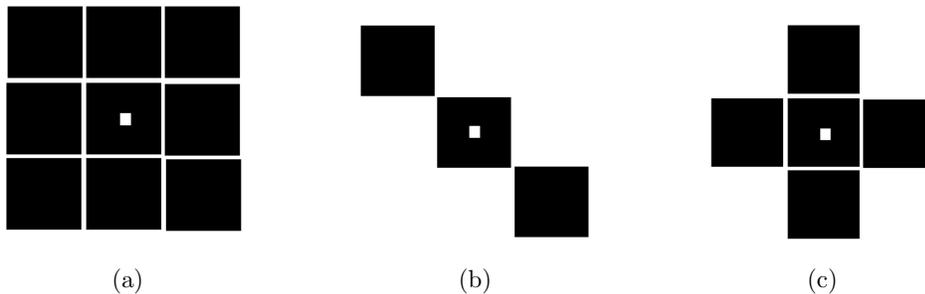


Figura 3.6: Exemplos de funções estruturantes planares contidos em uma janela 3×3 . O ponto branco denota o centro do elemento estruturante.

No caso de imagens em níveis de cinza, a grande maioria das aplicações considera apenas funções estruturantes planares. Entretanto, as funções estruturantes não-planares podem ser utilizadas na definição de transformações morfológicas dependentes de escala, as quais permitem a representação multi-escala do sinal analisado.

Um exemplo típico do emprego dessa representação é a análise de uma árvore. Em se tratando de um galho, uma distância variando de alguns centímetros a alguns metros é

mais apropriada, entretanto, para analisar uma floresta é mais adequado considerar uma medida de quilômetros, por exemplo.

Uma classe importante de tais funções está relacionada com a representação de espaços-escala morfológicos denotada aqui pela notação g_σ e definida seguindo Jackway e Deriche [1996]:

$$g_\sigma(a, b) = |\sigma| g(\sigma^{-1}a, \sigma^{-1}b) \quad a, b \in \mathcal{G}, \forall \sigma \neq 0, \quad (3.7)$$

em que σ transmite a noção de escala.

É importante observar ainda que, para garantir um comportamento espaço-escala condizente, a função g_σ deve satisfazer às seguintes condições [Jackway e Deriche, 1996]:

$$|\sigma| \rightarrow 0 \Rightarrow g_\sigma(a, b) = \begin{cases} 0, & \text{se } a = 0 \text{ e } b = 0, \\ -\infty, & \text{caso contrário.} \end{cases} \quad (3.8)$$

$$0 < |\sigma_1| < |\sigma_2| \Rightarrow g_{\sigma_1}(a, b) \leq g_{\sigma_2}(a, b) \quad \forall a, b \in \mathcal{G}_{\sigma_1}$$

$$|\sigma| \rightarrow \infty \Rightarrow g_\sigma(a, b) \rightarrow 0.$$

3.2.3 Operações básicas

A *erosão* e a *dilatação* constituem as transformações elementares, empregadas na definição de operações morfológicas mais complexas [Soille, 1999].

Sejam $f : \mathcal{D} \subset \mathbb{Z}^2 \rightarrow \mathbb{Z}$ uma função que define uma imagem e $g : \mathcal{G} \subset \mathbb{Z}^2 \rightarrow \mathbb{Z}$ uma função estruturante. As operações de erosão e dilatação podem ser definidas da seguinte forma:

Definição 3.1 *Erosão:* A erosão da função $f(x, y)$ por uma função estruturante $g(a, b)$, $[\epsilon_g(f)](x, y)$, é dada por:

$$[\epsilon_g(f)](x, y) = \min_{a, b \in \mathcal{G}} \{f(x + a, y + b) - g(a, b)\}. \quad (3.9)$$

Definição 3.2 *Dilatação:* A dilatação da função $f(x, y)$ por uma função estruturante $g(a, b)$, $[\delta_g(f)](x, y)$, é dada por:

$$[\delta_g(f)](x, y) = \max_{a, b \in \mathcal{G}} \{f(x + a, y + b) + g(a, b)\}. \quad (3.10)$$

As principais propriedades dessas transformações são resumidas a seguir, lembrando que a dilatação e a erosão são representadas por δ e ϵ , respectivamente.

- Comutatividade: $\delta_g(f) = \delta_f(g)$;
- Não-comutatividade: $\varepsilon_g(f) \neq \varepsilon_f(g)$;
- Associatividade: $\delta_{\delta_h(g)}(f) = \delta_h(\delta_g(f))$;
- Não-associatividade: No entanto, a seguinte relação é verdadeira: $\varepsilon_h(\varepsilon_g(f)) = \varepsilon_{\varepsilon_h}(f)$;
- Extensividade: $\delta_g(f) \subseteq f$ se o elemento estruturante contém sua origem;
- Anti-extensividade: $\varepsilon_g(f) \subseteq f$ se o elemento estruturante contém sua origem;
- Decomposição da dilatação: $\delta_{g \cup h}(f) = \delta_g(f) \cup \delta_h(f)$ e $\delta_{g \cap h}(f) = \delta_g(f) \cap \delta_h(f)$
- Decomposição da erosão: $\varepsilon_{g \cup h}(f) = \varepsilon_g(f) \cup \varepsilon_h(f)$ e $\varepsilon_{g \cap h}(f) = \varepsilon_g(f) \cap \varepsilon_h(f)$

Juntamente com a associatividade, as propriedades relacionadas à decomposição permitem a definição de formas de implementação eficientes dessas transformações, explorando-se, por exemplo, o conceito de separabilidade dos elementos estruturantes. Além disso, pode-se verificar que ambas as transformações são duais, ou seja,

$$\delta_g(f) = [\varepsilon_g(f^c)]^c \quad (3.11)$$

em que c denota o complemento. Isso justifica a relação de dualidade entre as propriedades apresentadas anteriormente.

Algumas dessas propriedades podem ser visualizadas na Figura 3.7, que apresenta um exemplo de erosão e dilatação para uma imagem em níveis de cinza. Observe que enquanto a erosão amplia as regiões escuras da imagem e reduz as claras, a dilatação faz o contrário em razão da dualidade (anti-extensividade e extensividade).

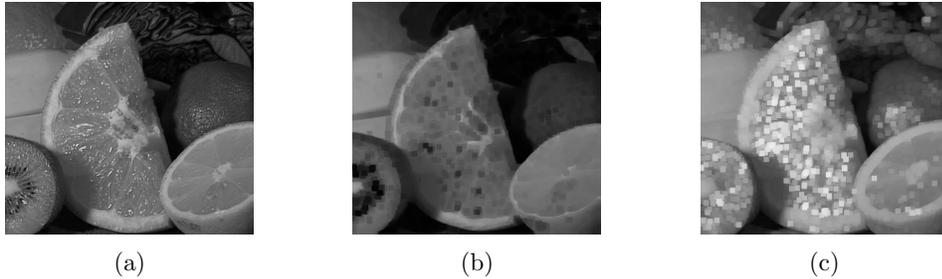


Figura 3.7: Exemplos de erosão e dilatação: (a) imagem original, (b) erosão considerando uma função estruturante planar 10×10 e (d) dilatação considerando esta mesma função.

Na próxima seção, outras transformações morfológicas são discutidas, incluindo os operadores de mapeamento do tipo *toggle*.

3.2.4 Outras operações

As transformações apresentadas na seção anterior possuem algumas limitações, no entanto, se combinadas adequadamente algumas interessantes propriedades podem ser observadas.

Considere, por exemplo, a erosão, que ao ser aplicada em uma imagem f utilizando um elemento estruturante g elimina todas as suas pequenas estruturas, menores que o elemento estruturante utilizado, eliminando também todas as demais estruturas. Todavia, se uma dilatação for realizada, subsequencialmente, partes das estruturas reduzidas serão recuperadas.

As operações de *abertura* e *fechamento* com uma função estruturante g , denotadas respectivamente por γ_g e ϕ_g , são obtidas combinando-se a erosão e a dilatação da seguinte forma:

$$\gamma_g(f) = \delta_{\tilde{g}}(\varepsilon(f)), \quad (3.12)$$

$$\phi_g(f) = \varepsilon_{\tilde{g}}(\delta(f)), \quad (3.13)$$

em que \tilde{g} corresponde ao transposto de g , ou seja, ao simétrico da função g em relação à sua origem.

De forma geral, estas transformações tendem a recuperar aproximadamente certas estruturas da imagens, removidas pela primeira transformação.

Assim, em termos visuais, a abertura regulariza os contornos e elimina pequenas “ilhas” e “cabos” estreitos de uma imagem binária, enquanto o fechamento suprime pequenos “canais” e “lagos” estreitos. Da mesma forma, para imagens em tons de cinza, a abertura elimina estruturas claras e o fechamento atua sobre estruturas escuras. A Figura 3.8 ilustra essas noções intuitivas. Sendo que a Figura 3.8(b) corresponde ao resultado da abertura utilizando um elemento estruturante circular de raio $r = 20$ e a Figura 3.8(c) apresenta o resultado do fechamento realizado utilizando um elemento estruturante circular de raio $r = 5$.

É importante observar que essas transformações são duais, crescentes e idempotentes. Transformações crescentes e idempotentes caracterizam os denominados *filtros morfológicos*, uma abordagem não-linear que conduz tanto à recuperação de imagens com ruído quanto a remoção de estruturas ou objetos específicos de uma imagem seguindo critérios, tais como geometria ou contraste. Por se tratar de uma transformação crescente, a relação de ordem entre as imagens filtradas é preservada. A idempotência, por sua vez, garante a estabilidade da operação, evitando problemas como oscilação dos resultados, por exemplo.

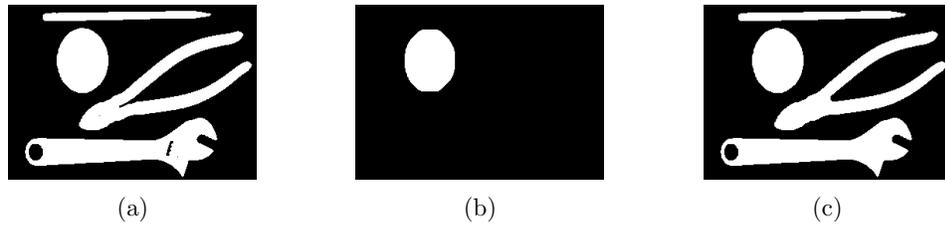


Figura 3.8: Exemplos de abertura e fechamento: (a) imagem original, (b) resultado da abertura utilizando um elemento estruturante circular de raio $r = 20$ e (c) fechamento utilizando um elemento estruturante circular de raio $r = 5$.

A Figura 3.9 ilustra a aplicação das operações de abertura e fechamento em uma imagem com ruído do tipo sal-e-pimenta. Observe que, como nenhuma dessas transformações satisfazem à propriedade de auto-dualidade, não há um tratamento simultâneo das estruturas claras e escuras da imagem. Este problema pode ser minimizado combinando as duas diferentes transformações.

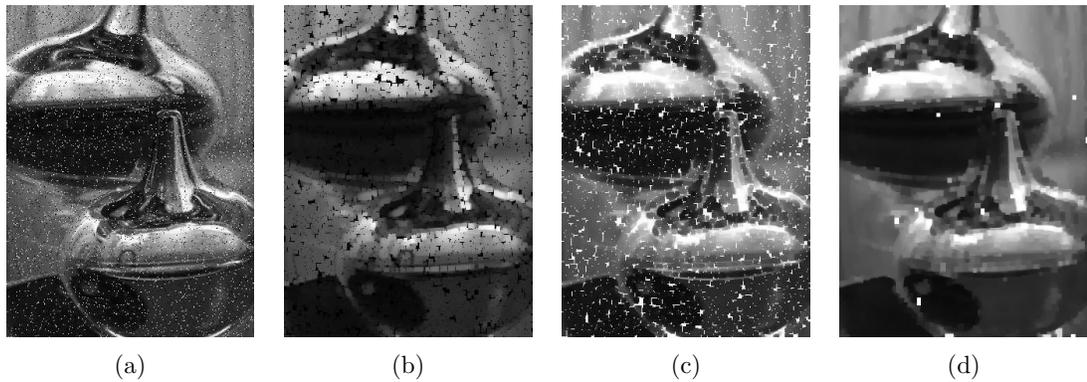


Figura 3.9: Exemplos de filtragem morfológica de uma imagem com ruído do tipo sal-e-pimenta: (a) imagem original com ruído, (b) resultado da abertura utilizando uma função estruturante circular de raio $r = 3$, (c) imagem obtida após o fechamento, utilizando essa mesma função estruturante e (d) resultado da aplicação de um fechamento seguido de uma abertura.

Observe que a composição sucessiva dessas transformações também constitui um filtro morfológico. Assim, dependendo do tipo de ruído, pode-se combiná-las de forma a produzir resultados mais adaptados a cada situação. Tal combinação consiste nos chamados filtros alternados sequenciais [Soille, 1999].

Finalmente, a definição a seguir introduz uma categoria de transformações de mapeamento para todos os pontos da imagem denominada *toggle*. Em transformações do tipo *toggle*, a ideia central consiste em associar uma imagem com uma série de primitivas φ_i , a

partir de uma regra de decisão que determina qual dessas primitivas deve ser considerada em cada ponto da imagem [Soille, 1999, Serra et al., 1992].

Definição 3.3 *Denomina-se mapeamento toggle das primitivas φ_i qualquer mapeamento w tal que [Serra et al., 1992]:*

- em cada ponto (x, y) , $w(x, y)$ é igual a φ_i ou $f(x, y)$,
- no mapeamento w , o critério que transforma uma das primitivas φ_i em um dado ponto (x, y) depende somente das primitivas φ_i , do valor numérico $f(x, y)$ e de possíveis constantes.
- se no ponto (x, y) ao menos uma das primitivas φ_i coincide com o mapeamento identidade $f(x, y)$, então $w(x, y) = f(x, y)$

Um exemplo simples de mapeamento toggle é a limiarização binária, em que as primitivas consistem em duas imagens branca e preta e a regra de decisão envolve, no ponto (x, y) , o valor de $f(x, y)$ e uma constante denominada limiar.

Outro exemplo de operador de mapeamento consiste no conhecido *toggle* de contraste k^2 definido da seguinte forma [Soille, 1999]:

$$k^2(x, y) = \begin{cases} \varphi_1(x, y), & \text{se } \varphi_1(x, y) - f(x, y) < \varphi_2(x, y) - f(x, y), \\ \varphi_2(x, y), & \text{caso contrário.} \end{cases} \quad (3.14)$$

em que φ_1 e φ_2 indicam duas transformações duais, extensiva e anti-extensiva, respectivamente. A regra de decisão envolve os valores da imagem processada f e das duas transformações φ_1 e φ_2 . Em outras palavras, o valor de k^2 , em cada ponto (x, y) , é igual ao valor da transformação mais próxima ao valor da imagem f nesse ponto. Usualmente, nesse caso, as transformações φ_1 e φ_2 correspondem à dilatação e à erosão, respectivamente, considerando uma função estruturante planar.

Neste trabalho, as propriedades dessa classe de transformação foram exploradas na definição de um operador de mapeamento utilizado para o cálculo de uma estimativa do background das imagens de impressões digitais. Tal mapeamento é apresentado no próximo capítulo e se mostrou bastante eficaz na segmentação de imagens provenientes de diferentes sensores.

3.3 Operadores direcionais

Uma importante característica para a descrição e/ou reconhecimento de objetos em imagens digitais é a orientação de suas estruturas, representando, por exemplo, aspectos

interessantes de forma e textura. No contexto de impressões digitais, as informações direcionais são definidas a partir do *campo direcional*.

O *campo direcional* fornece uma visão global da impressão digital, descrevendo localmente as direções de cada um dos pontos que compõem a imagem. Através dele é possível detectar, por exemplo, transições entre diferentes direções e, assim, extrair *pontos singulares* [Bazen e Gerez, 2002], ou, ainda, encontrar regiões de quebras com descontinuidade na representação das papilas e conectá-las [Oliveira e Leite, 2008].

As subseções seguintes ilustram dois operadores direcionais a serem considerados posteriormente neste trabalho. O primeiro deles baseia-se na informação obtida pelo gradiente [Bazen et al., 2004, Jin e Kim, 2010] e é descrito na Subseção 3.3.1. A Subseção 3.3.2 apresenta o operador proposto por Oliveira e Leite [2008] que pode ser visto como uma generalização dos métodos baseados em projeções direcionais discretas [Mehtre, 1993, Maltoni et al., 2009].

Informações adicionais e outros operadores podem ser encontrados, por exemplo, em Maltoni et al. [2009], Terol-Villalobos et al. [2008] e Gottschlich et al. [2009].

3.3.1 Método do gradiente

A abordagem mais popular encontrada na literatura para estimativa do *campo direcional* foi introduzida por Kass e Witkin [1987], e é comumente conhecida como *método do gradiente*. A principal vantagem desta abordagem em relação às demais é que ela não se limita a um número fixo de direções.

Em resumo, a informação direcional de um determinado ponto da imagem corresponde simplesmente à direção perpendicular ao vetor de gradientes horizontal e vertical nesse ponto. No entanto, como essa informação é obtida considerando uma escala muito fina, um processo de filtragem deve ser realizado a fim de reduzir componentes de ruído.

Esta filtragem deve ser realizada com cautela, pois os vetores de gradiente não podem ser diretamente filtrados, uma vez que vetores opostos, apesar de indicarem uma mesma direção, podem se anular.

Com o objetivo de evitar esse problema, Rao e Jain [1992] propuseram um método, chamado de *gradiente quadrático*, no qual os vetores de gradientes são multiplicados para que, assim, vetores opostos representem uma mesma direção. O vetor com o *gradiente quadrático* é expresso de acordo com

$$[G_{sx}, G_{sy}] = [(G_x^2 - G_y^2), (2G_x G_y)], \quad (3.15)$$

em que G_x e G_y referem-se às derivadas da imagem em relação a direção horizontal e vertical, respectivamente. Assim, a imagem com o campo direcional, $d : \mathcal{D} \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, é calculada através de

$$d(x, y) = \frac{\pi}{2} + \frac{1}{2} \tan^{-1} \left(\frac{\sum_W G_{sx}}{\sum_W G_{sy}} \right). \quad (3.16)$$

G_{sx} e G_{sy} indicam as componentes horizontal e vertical do vetor de gradiente quadrático, respectivamente, e W é uma janela *gaussiana* de tamanho arbitrário centrada no ponto (x, y) .

Note que o tamanho dessa janela pode ser visto como uma representação da escala do campo direcional nos termos da representação do espaço-escala linear discutida em Witkin [1984], Lindeberg [1994] e Koenderink [1984]. Sendo que quanto maior a mesma mais grosseira é a escala [Bazen et al., 2004, Jin e Kim, 2010]. Normalmente as derivadas são computadas pela convolução da imagem com uma máscara de Sobel [Gonzalez e Woods, 2001] ou outras máscaras de gradiente [Ando, 2000].

3.3.2 Operador multi-escala direcional

O operador multi-escala direcional pode ser visto como uma generalização do método para estimativa do campo direcional apresentado por Mehtre [1993], dentre outros. Ambos baseiam-se na observação de que, geralmente, em imagens de impressões digitais o contraste obtido entre a informação seguindo a direção da papila (cristas e vales) e sua direção perpendicular é maior do que aquele obtido considerando-se outros pares de direções perpendiculares.

A Figura 3.10 ilustra essa observação. Para tanto, considere as barras no centro da figura. A barra vertical corresponde à informação obtida na direção da papila enquanto a barra horizontal, à informação considerando sua direção perpendicular.

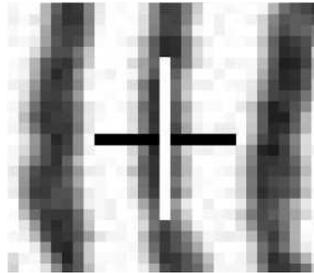


Figura 3.10: Detalhe do contraste obtido entre a informação na direção das papilas e sua direção perpendicular.

Informalmente, o operador estima a orientação de cada ponto (x, y) da imagem, dividindo um semicírculo $([0^\circ, 180^\circ])$ em D direções discretas $([0, D - 1])$, e considerando o

contraste obtido entre os desvios-padrão calculados para os valores da imagem ao longo de um conjunto de segmentos de linhas direcionais.

As coordenadas (x, y) dos pontos em um segmento de linha discreto com tamanho n e direção α são calculadas considerando-se uma janela deslizante Γ de tamanho $n \times n$, de tal forma que

$$\begin{aligned} x &= x_{centro} + p \times \cos(\alpha) \\ y &= y_{centro} + p \times \sin(\alpha), \end{aligned} \quad (3.17)$$

para todo p discreto, tal que $-n/2 \leq p \leq n/2$. x_{centro} e y_{centro} são as coordenadas dos pontos que contêm a janela deslizante Γ centrada nessa posição.

Observe que tais pontos podem ser entendidos como os vizinhos (ou *vértices adjacentes*, na representação em grafo) do ponto (x_{centro}, y_{centro}) considerando uma determinada direção (α) e escala (n). A Figura 3.11 mostra os pontos que compõem um segmento de linha de direção $\alpha = 45^\circ$ e tamanho $n = 9$, computado através da Equação 3.17.

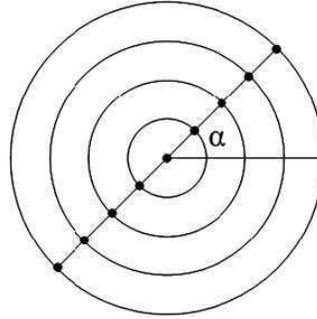


Figura 3.11: Exemplo de pontos contidos em segmento de linha de direção igual a 45° e tamanho 9.

O conjunto s_i^n com D segmentos de linha discretos de tamanho n e direção discreta i , é calculado repetindo o procedimento descrito na Equação 3.17 para todas as D direções ($i \in \{0, 1, \dots, D-1\}$), alterando o valor de α apropriadamente, ou seja, $\alpha = 0, 1 \times 180/D, 2 \times 180/D, \dots, (D-1) \times 180/D$.

Note que para uma janela bidimensional de dimensão $n \times n$ é possível representar $(2 \times n - 1)$ direções [Soille, 1999].

A versão mais detalhada da imagem direcional, $d' : \mathcal{D} \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, é então calculada de acordo com a seguinte equação:

$$d'(x, y) = \begin{cases} i, & \text{se } std(s_i^n) < std(s_{\perp(i)}^n) \\ \perp(i), & \text{se } std(s_i^n) > std(s_{\perp(i)}^n) \\ v, & \text{caso contrário,} \end{cases} \quad (3.18)$$

em que std indica o desvio padrão e i e $\perp(i)$ correspondem ao par de direções perpendiculares que possuem contraste máximo ou mais formalmente, equivalem ao par k e $\perp(k)$ obtido em

$$\arg\text{-max}_{k \in \{0, \dots, D/2-1\}} \{std(s_k^n) - std(\perp(s_k^n))\}. \quad (3.19)$$

v é um valor distinguível (o valor atribuído para v neste trabalho foi 255) que representa um região homogênea onde não há direção dominante.

Finalmente, a imagem contendo o campo direcional, $d : \mathcal{D} \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, é obtida considerando uma janela $\Omega : \mathcal{D} \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, também conhecida como *janela de filtragem*, centrada em cada pixel da imagem d' , conforme a seguinte equação:

$$d(x, y) = mode_{a, b \in \mathcal{W}} \{d'(x + a, y + b)\} \quad (3.20)$$

em que \mathcal{W} indica o domínio da janela de filtragem e *mode*, o operador estatístico moda que, nesse caso, computa a direção mais frequente em cada janela.

Note que o tamanho desta *janela de filtragem* constitui um fator de escala do campo direcional, de tal modo que uma janela pequena produz uma representação *fina* (com mais detalhes) do campo direcional correspondente, enquanto uma janela maior produz uma representação *grosseira* (com poucos detalhes).

Essa característica pode ser explorada de diferentes formas. Neste trabalho, por exemplo, escalas maiores foram empregadas no delineamento da área da impressão digital a partir da consideração apenas das informações direcionais dominantes, como será visto no capítulo seguinte que descreve o método proposto.

Capítulo 4

Método Proposto

Este capítulo descreve um novo método *pontual* de segmentação de imagens de impressões digitais que explora as propriedades de simplificação de um operador do tipo *toggle* e do operador multi-escala direcional introduzido na Seção 3.3.2. O método aqui definido possui uma interessante característica que o diferencia das principais abordagens de segmentação encontradas na literatura. Mais especificamente, não requer configurações ou treinamentos específicos para uma dada classe de imagens ou sensor, o que diz respeito diretamente ao problema da interoperabilidade mencionado no Capítulo 2.

A Seção 4.1, a seguir, fornece uma descrição geral do método proposto e a Seção 4.2 apresenta algumas propriedades do operador *toggle* proposto relacionadas ao contraste da imagem. A Seção 4.3 discute a relação dos parâmetros deste operador com as estruturas da imagem. Na Seção 4.4 são apresentadas algumas considerações sobre a teoria espaço-escala e o operador proposto. A Seção 4.5 descreve uma extensão para o método definido neste capítulo, utilizando as transformações propostas para modificar o método de Bazen e Gerez [2001].

4.1 Descrição geral do método

De uma maneira geral, o método de segmentação proposto neste trabalho consiste dos seguintes passos apresentados no fluxograma da Figura 4.1.

Primeiramente, o bloco *Detecção fina do campo direcional* estima a orientação da *imagem de entrada* considerando o operador multi-escala direcional (Subseção 3.3.2) utilizando uma *janela de filtragem* pequena, Ω_s , de tamanho 5×5 .

Esta operação de filtragem define uma imagem de orientação, w , que corresponde a uma representação em escala fina do campo direcional. Tal representação preserva detalhes direcionais importantes que são usados no passo seguinte, denominado *subtração do fundo*, que realiza a subtração do *background* da imagem original.

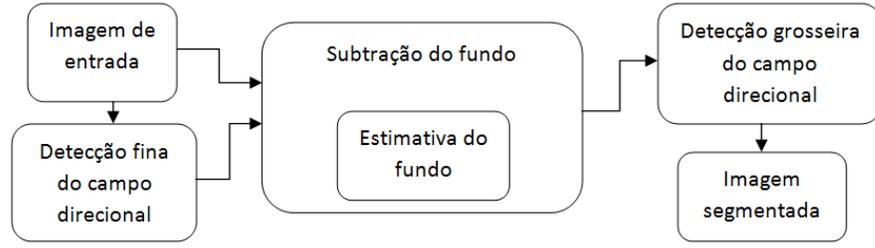


Figura 4.1: Fluxograma do método proposto.

Esta operação de subtração separa a região de interesse, contendo cristas e vales, do restante da imagem, da seguinte forma:

$$f'(x, y) = f(x, y) - b(x, y), \quad (4.1)$$

em que f' representa a imagem com a impressão digital contendo a informação das papilas, f indica a imagem de entrada original e b , uma estimativa para seu *background*.

Para obter uma estimativa deste background, foi definido um mapeamento que explora algumas propriedades das transformações morfológicas dependentes de escala (subseção 3.2.2). Essas propriedades incluem, por exemplo, a supressão de extremos da imagem (mínimos e máximos regionais) e sua simplificação.

O mapeamento *toggle* correspondendo à estimativa do background b é dado por:

$$b(x, y) = \begin{cases} f(x, y), & \text{se } \phi_1^k(x, y) - f(x, y) \leq f(x, y) - \phi_2^k(x, y), \\ \phi_3(x, y), & \text{caso contrário,} \end{cases} \quad (4.2)$$

em que, novamente, f corresponde à imagem de entrada, $\phi_1^k = [\delta_{g_\sigma}(f)]^k$ e $\phi_2^k = [\epsilon_{g_\sigma}(f)]^k$ são, respectivamente, a dilatação e a erosão de f com uma função estruturante não-planar g_σ , k vezes, e $\phi_3 = [\delta_{g_\perp w}(f)]$ indica a dilatação linear de f na direção perpendicular à orientação contida na imagem de representação fina do campo direcional, w , localizada no ponto (x, y) .

Informalmente, o mapeamento definido na Equação 4.2 define, como estimativa para b , o próprio valor da imagem original de entrada, f , nos pontos que convergem para uma região de máximos localizados em áreas uniformes ou quase uniformes, baseando-se na proximidade de $f(x, y)$ com a dilatação $\phi_1^k(x, y)$, e um valor que preservará a informação papilar, representado pela dilatação perpendicular $\phi_3(x, y)$ que leva em conta a orientação das cristas na imagem original f .

As Figuras 4.2 e 4.3 ilustram as transformações descritas acima para imagens ruidosas. A Figura 4.2 corresponde ao passo de *subtração de fundo* definido pela Equação 4.1. Vale a pena notar que as regiões correspondendo ao fundo da imagem e aos vales (sem

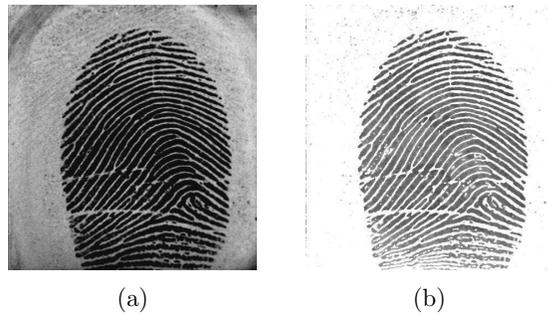


Figura 4.2: Exemplo de segmentação para uma imagem do FVC2000: (a) imagem original e (b) resultado da aplicação da equação 4.1.

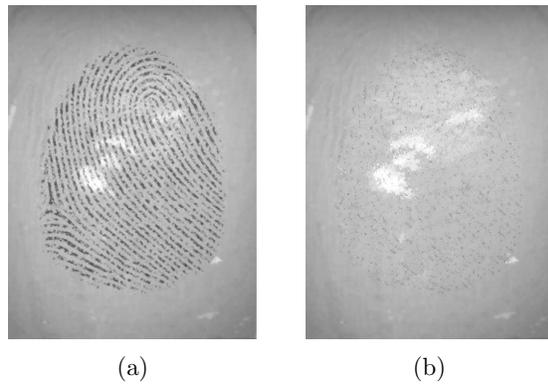


Figura 4.3: Estimativa de background computada utilizando $k = 15$ e $\sigma = 20$.

informações papilares) convergem para um máximo regional representado aqui pelas cores mais claras da Figura 4.2 (b). A Figura 4.3, por sua vez, exemplifica a estimativa do background definida pela Equação 4.2.

Finalmente, o bloco *Detecção grosseira do campo direcional*, apresentado na Figura 4.1, considera o operador multi-escala direcional e uma janela de filtragem grande, Ω_l , de tamanho 45×45 , a fim de regionalmente definir a área dominante da imagem contendo as informações papilares de f' . Nesta representação grosseira da informação direcional, as linhas externas das regiões de mesma orientação constituem o resultado da segmentação.

A Figura 4.4 ilustra todos os passos do método descritos neste capítulo. A Figura 4.4 (b) apresenta o resultado do passo de *Detecção fina do campo direcional*, enquanto as Figuras 4.4 (c) e (d) mostram o resultado da *Estimativa do fundo* e da *Subtração de fundo* definidas, respectivamente, pelas Equações 4.2 e 4.1. As Figuras 4.4 (d) e (f) apresentam o resultado do passo de *Detecção grosseira do campo direcional* e a imagem final segmentada.

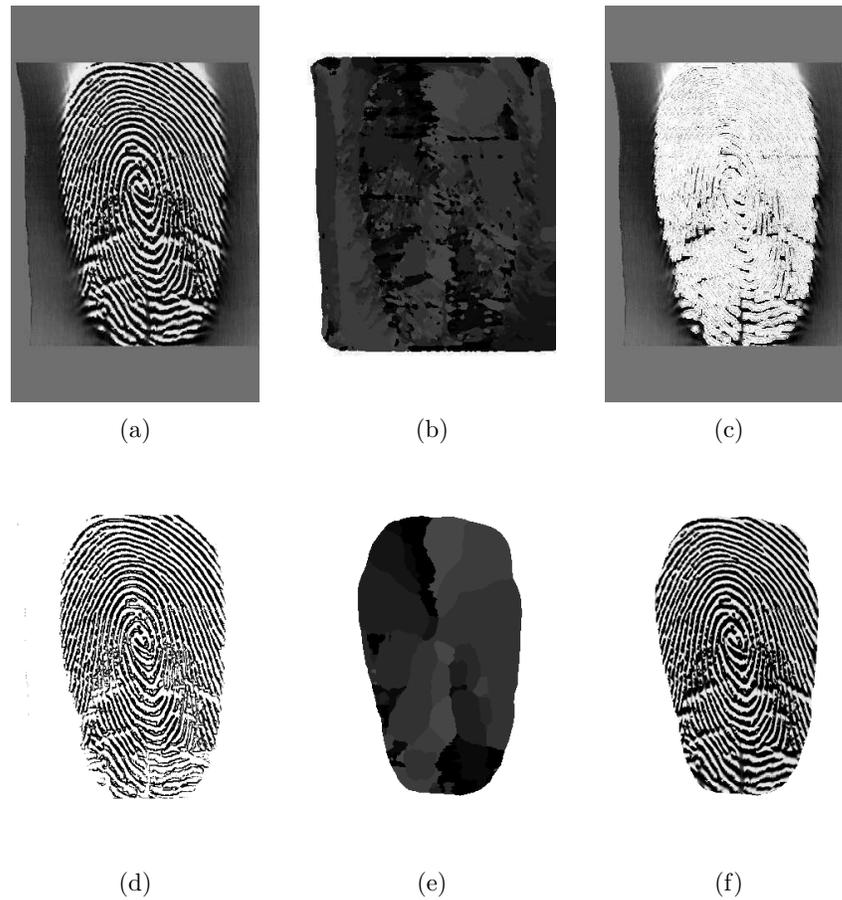


Figura 4.4: Exemplo dos passos realizados na segmentação de uma imagem do FVC2004: (a) imagem original, (b) orientação fina do campo direcional, (c) estimativa do background (d) resultado do processo de subtração do fundo, (e) orientação grosseira do campo direcional computada considerando a imagem de subtração do fundo (nesta figura, cada um dos tons de cinza indica uma direção diferente) e (f) imagem final segmentada.



Figura 4.5: Relação de proximidade entre os valores transformados e a imagem original: (a) imagem original, (b) mapa de proximidade considerando a escala $\sigma = 10$ e número de iterações $k = 10$, (c) mapa de proximidade considerando o mesmo número de iterações e a escala $\sigma = 20$ e (d) mapa de proximidade considerando a escala $\sigma = 30$ e o mesmo número de iterações.

4.2 Aspectos relacionados ao contraste

Como mencionado, o mapeamento apresentado aqui e utilizado para estimativa do background da imagem explora algumas propriedades das transformações morfológicas primitivas (erosão e dilatação) dependentes de escala (Subseções 3.2.2 e 3.2.3), bem como a proximidade dos valores transformados por elas com os valores correspondentes na imagem original. Por exemplo, nessas transformações quanto maior a escala, maior deve ser a diferença entre os níveis de cinza para que haja uma diferença entre os valores transformados. Essa propriedade pode ser observada na Figura 4.5 que ilustra também a relação de proximidade entre os valores transformados e a imagem original dependendo da escala escolhida.

Nesta figura, os pontos cinzas correspondem a regiões onde não há diferença entre os valores transformados pelas primitivas em uma dada escala. Os pontos brancos, por sua vez, indicam regiões em que o valor do dilatado está mais próximo do seu valor correspondente na imagem original, e os pontos pretos, em contrapartida, representam regiões em que o valor do erodido está mais próximo.

Observe que, a medida que a escala aumenta, regiões menores são afetadas pelas primitivas. Isso acontece porque só há diferença entre os valores transformados se há também uma certa diferença entre os níveis de cinza. Tal diferença deve ser proporcional à altura da função estruturante utilizada, ou seja, proporcional à escala. Dessa maneira, pode-se utilizar uma escala maior para considerar somente regiões de alto contraste.

Essa propriedade de filtragem é a base do operador de mapeamento proposto neste trabalho e foi utilizada na definição da sua *regra de decisão*. Como apresentado anteriormente, essa regra de decisão avalia a proximidade do valor dilatado em um determinado pixel, com o valor correspondente na imagem original, bem como a proximidade dos va-

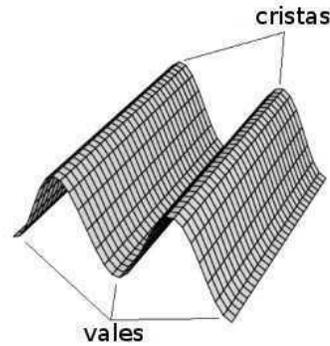


Figura 4.6: Representação 3D de uma impressão digital ideal.

lores das duas transformações (dilatação e erosão) para, então, decidir se a informação direcional deve ser considerada.

Isso acontece porque as imagens de impressão digital são compostas por vários segmentos de linha claros e escuros denominados vales e cristas, respectivamente, que definem regiões de alto contraste, as quais constituem diferenças nos valores das transformações. Esses segmentos de linha configuram o padrão ideal de listras ilustrado em 3D na Figura 4.6.

A Figura 4.7 mostra o resultado dessas transformações para uma imagem de impressão digital a partir de diferentes escalas. Assim como na Figura 4.5, nesta figura, os pontos brancos e pretos representam, respectivamente, regiões em que o valor da dilatação ou da erosão está mais próximo do valor correspondente na imagem original e os pontos cinzas representam regiões em que não há diferença entre esses valores. Observe que, à medida que a escala aumenta, as regiões afetadas pelas transformações se concentram em regiões de mais alto contraste. Note ainda que essa região se regulariza com a forma da impressão quando a escala $\sigma = 20$ é considerada.

Nesta figura pode-se notar ainda uma relação entre as regiões de mais alto contraste e a escala, σ , utilizada. Esse aspecto pode ser explorado, por exemplo, em aplicações que envolvem a avaliação da qualidade de imagens de impressões digitais.

A Figura 4.8 ilustra a importância de se considerar funções estruturantes não-planares na Equação 4.2. Nesta figura, observa-se claramente a não convergência das imagens correspondentes à estimativa do background. Isto acontece porque, no limite, a aplicação deste mapeamento com funções não-planares equivale a uma limirização simples, em que atribui-se branco aos pontos mais próximos do menor valor em tons de cinza da imagem e preto aos demais pontos.

De acordo com a Equação 4.2, como apenas os pontos da imagem original próximos da erosão são transformados, somente as informações que se referem às cristas da imagem

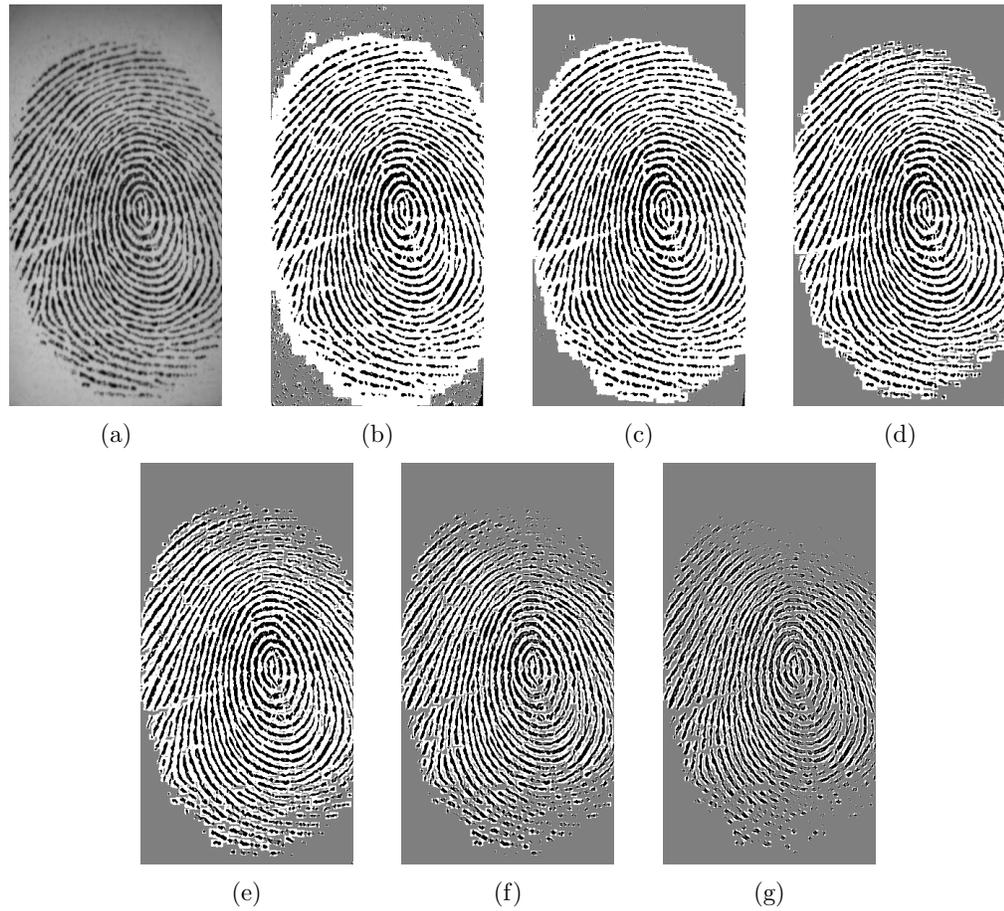


Figura 4.7: Relação de proximidade para uma imagem de impressão digital: (a) imagem original, mapa de proximidade considerando o número de iterações $k = 15$ e escalas (b) $\sigma = 5$, (c) $\sigma = 10$, (d) $\sigma = 20$, (e) $\sigma = 30$, (f) $\sigma = 40$ e (g) $\sigma = 50$.

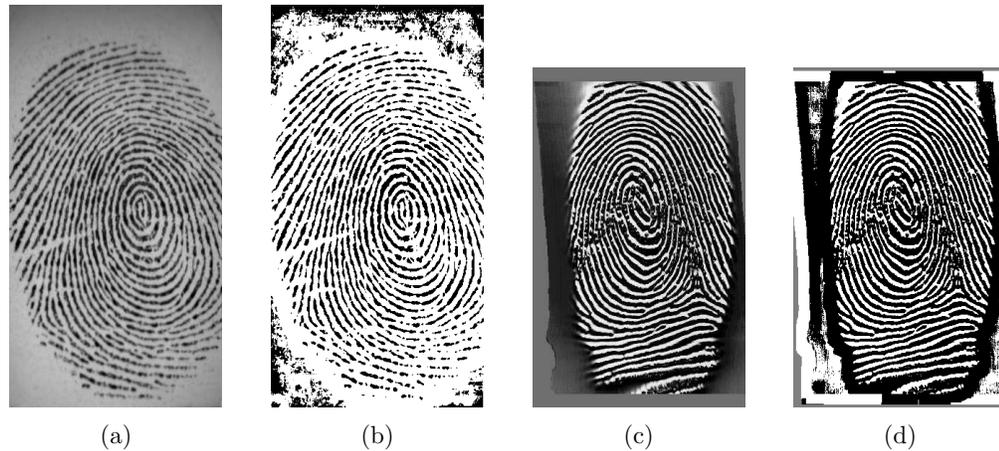


Figura 4.8: Mapa de proximidade utilizando elementos estruturantes planares: (a) imagem original, (b) mapa de proximidade com $k = 15$, (c) imagem original e (d) mapa de proximidade com o mesmo número de iterações k .

são afetadas pela transformação do mapeamento correspondente à dilatação ϕ_3 .

Como visto anteriormente, esta transformação calcula uma dilatação linear realizada na direção perpendicular à orientação das cristas, com o objetivo de garantir que regiões onde há borramento não sejam preservadas durante a subtração do background. Isso acontece porque em regiões de borramento não se configura aproximadamente o padrão mostrado na Figura 4.6, e dessa forma, o valor da dilatação, ϕ_3 , se aproxima do valor original da imagem e a subtração, f' , se aproxima de zero. Assim, a diferença entre o valor transformado por ϕ_3 , num determinado ponto, e o valor correspondente na imagem original é máximo somente em regiões onde há contraste direcional.

A Figura 4.9 ilustra o resultado da aplicação da equação de subtração de fundo para uma imagem que contém uma região com borramento. Pode-se notar que o valor do mapeamento definido pela Equação 4.2, na Figura 4.9 (a), em tal região, converge para um valor próximo do seu correspondente na imagem original e, assim, a região de borramento é desconsiderada na segmentação, como mostra a Figura 4.9 (b).

4.3 Parâmetros versus estruturas da imagem

O método de segmentação descrito neste capítulo depende de três parâmetros relacionados ao mapeamento definido pela Equação 4.2. Tais parâmetros referem-se à escala, σ , e ao número de iterações, k , empregados nas transformações dependentes de escala, ϕ_1 e ϕ_2 , e ao tamanho do elemento estruturante linear, $|g|$, empregado na dilatação ϕ_3 .

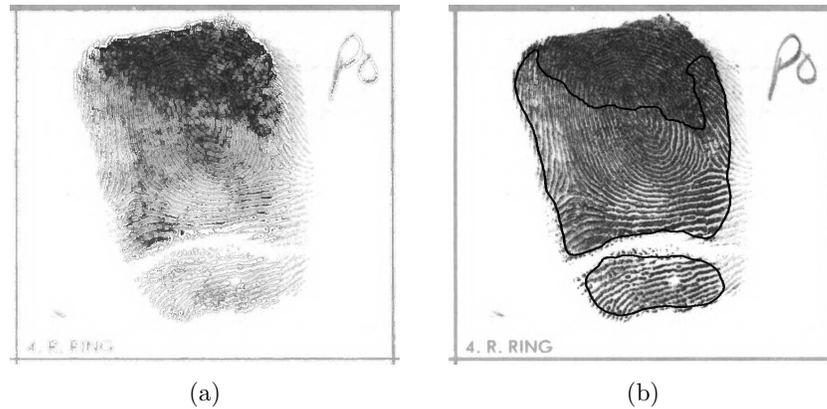


Figura 4.9: Exemplo de segmentação para uma imagem pertencente à base de dados do NIST com uma região de borramento: (a) corresponde à estimativa do fundo (Equação 4.2) e (b) ao resultado da segmentação.

Como discutido na seção anterior, a escala está relacionada ao contraste da imagem e, assim, pode ser definida para representá-lo. A Figura 4.7 apresenta um exemplo de aplicação do operador considerando diferentes escalas. Nela pode-se observar que, na escala $\sigma = 20$, a região transformada se aproxima da área da impressão digital.

Os outros dois parâmetros k e $|g|$, contudo, estão relacionados com o tamanho das estruturas da impressão digital (cristas e vales) nos termos discutidos a seguir.

O número de iterações, k , define o tamanho da vizinhança a ser considerada pelas transformações ϕ_1 e ϕ_2 . Mais precisamente, realizar k iterações de uma determinada transformação, utilizando um elemento estruturante 3×3 , é o mesmo que realizar uma iteração com um elemento estruturante $2k+1 \times 2k+1$. Assim, este parâmetro estabelece o tamanho da região de influência dos extremos (máximos e mínimos regionais) da imagem, ou seja, no contexto deste trabalho, esse valor define o tamanho das regiões de influência das cristas e vales da impressão digital.

A Figura 4.10 ilustra o resultado do mapeamento, b , definido pela Equação 4.2 quando diferentes números de iterações são considerados para uma mesma escala. Observe que, à medida que o número de iterações aumenta e se aproxima do tamanho das estruturas, o valor da imagem com a estimativa do background, b , nas regiões das papilas, aproxima-se do ideal.

Finalmente, pode-se notar que o parâmetro, $|g|$, que indica o tamanho da vizinhança a ser considerada pela dilatação ϕ_3 , também está relacionado com o tamanho das estruturas da impressão digital (distância entre cristas), pois define um limite para os borramentos na estrutura da imagem.



Figura 4.10: Exemplo de aplicação do mapeamento da Equação 4.2 para uma mesma escala $\sigma = 20$ e diferentes números de iterações: (a) imagem original, (b) resultado utilizando número de iterações $k = 1$, (c) $k = 5$, (d) $k = 10$ e (e) $k = 15$.

4.4 O operador toggle e a teoria espaço-escala

Neste trabalho, considerou-se uma função estruturante não-planar em forma de pirâmide, conforme ilustra a Figura 4.11, definida por $g(a, b) = \max\{|a|, |b|\}$ e cuja versão escalonada é dada por

$$g_\sigma(a, b) = -\sigma^{-1} \max\{|a|, |b|\}. \quad (4.3)$$

Especificamente, considera-se aqui uma função estruturante elementar, de tamanho 3×3 , em que g_σ é 0 na posição $(0, 0)$ e $-\sigma^{-1}$ em todas as outras posições. Como largamente explorado nos trabalhos de Jackway e Deriche [1996] e Dorini e Leite [2007], esta função estruturante não-planar pode ser empregada na definição de transformações morfológicas dependentes de escala, as quais permitem a análise de diferentes níveis de representação com interessantes propriedades de simplificação incluindo, por exemplo, a filtragem monotônica de extremos (máximos e mínimos) da imagem, ou seja, sem a criação de novos destes extremos.

Essas propriedades de simplificação constituem as condições necessárias para definição de transformações *espaços-escala*. Resumidamente, a teoria espaço-escala apresenta as condições necessárias para a definição de transformações que possibilitem a manipulação de características presentes em diferentes níveis de representação de maneira consistente.

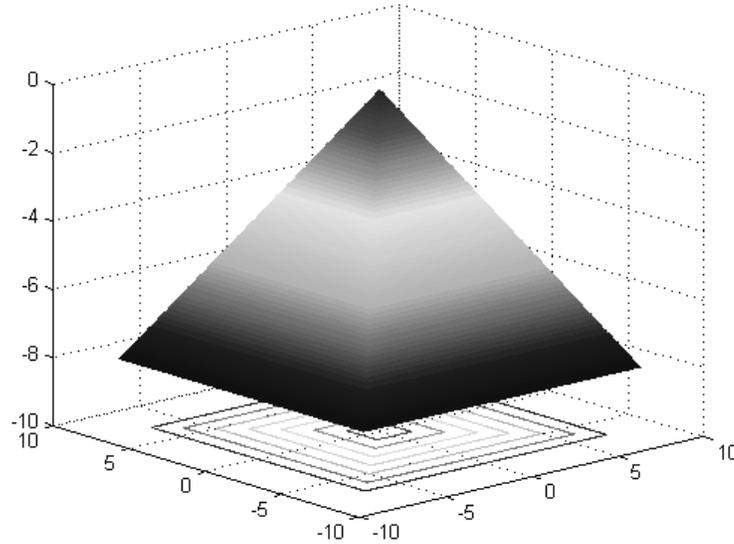


Figura 4.11: Função estruturante pirâmide.

Uma propriedade fundamental é que uma transformação para um nível mais grosso não introduza novas estruturas, ou seja, características pertencentes a uma escala específica também estão presentes em todas as escalas mais finas. Tal propriedade é denominada *monotonicidade*, já que o número de características descreve uma função monotonicamente crescente em função da escala.

Outra propriedade exigida para a abordagem espaço-escala é a *fidelidade*. Sejam $f(x, y)$ a imagem original e $T[x, y, \sigma](f)$ a representação da observação na escala σ . A fidelidade requer que a transformação espaço-escala convirja para o sinal original conforme σ tende a zero.

A *invariância euclidiana* também é desejada, ou seja, uma translação e/ou rotação da imagem original deve implicar em um espaço escala igualmente transladado/rotacionado.

Observe que os requisitos para a definição de um espaço-escala conduzem naturalmente a uma simplificação do sinal em consideração. Ao contrário de outras abordagens multi-escala, tais como pirâmides, a representação espaço-escala preserva a mesma resolução espacial (mesma amostragem) em todos os níveis. Isso permite o acesso imediato aos dados de interesse em qualquer escala sem a necessidade de processamento adicional. Aliado à propriedade de continuidade, esse fato permite rastrear de forma direta as características de interesse através da representação, bem como determinar sua localização exata no sinal original.

Assim, no contexto da teoria da morfologia matemática, desde que as determinadas condições sejam satisfeitas (subseção 3.2.2), as transformações dependentes de escala,

em geral, definem representações com propriedades espaço-escalas, como, por exemplo, o toggle auto-dual MMDE (*Self-dual Multi-scale Morphological Toggle*) definido em [Dorini e Leite, 2007, 2009].

Como visto anteriormente, algumas dessas propriedades foram consideradas experimentalmente, neste trabalho, na definição do operador representado pela Equação 4.2. Uma análise mais detalhada das propriedades deste operador e sua relação com a teoria espaço-escala precisa ser formalmente definida.

4.5 Extensões

Em resumo, a ideia principal do método apresentado neste capítulo consiste em uniformizar imagens com diferentes características de iluminação provenientes de diferentes bases de dados ou sensores, a partir de um conjunto de transformações morfológicas dependentes de escala. Dessa maneira, é possível realizar a segmentação sem a necessidade de configurações ou parâmetros específicos para cada base de dados ou sensor, o que, como mencionado anteriormente, refere-se ao problema de interoperabilidade de sensor.

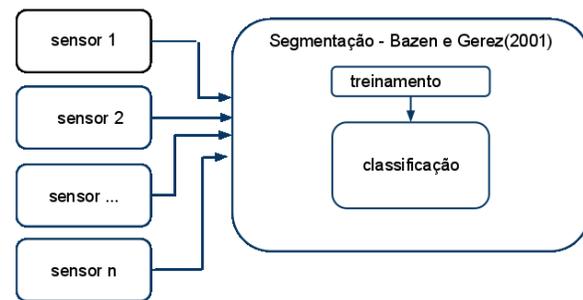
Nesse sentido, as transformações definidas aqui podem ser associadas a outros métodos de segmentação encontrados na literatura. O método definido por Bazen e Gerez [2001], por exemplo, pode ser modificado para considerar as imagens obtidas através do processo de subtração de background.

No método de Bazen e Gerez [2001] a base de dados com as imagens analisadas é dividida em dois conjuntos denominados *de treinamento* e *de classificação*. O *conjunto de treinamento* contém uma amostra representativa (usualmente aleatória) da base de dados que é utilizada pelo classificador no *treinamento* da rede neural, ou seja, na definição dos pesos que melhor classificam essas imagens, e o *conjunto de classificação* é composto por todas as imagens a serem classificadas.

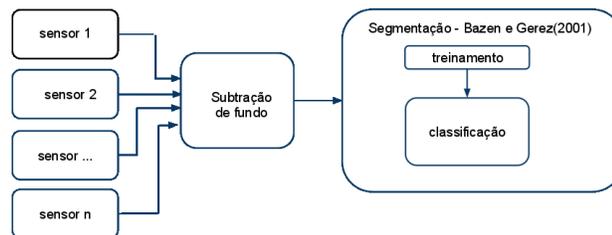
No caso da modificação proposta, tanto as imagens do conjunto de treinamento quanto as do conjunto de classificação devem ser processadas pelas Equações 4.1 e 4.2 antes de serem consideradas pelo classificador. Assim, imagens provenientes de diferentes sensores (ou bases de dados) podem ser utilizadas em um mesmo treinamento e os pesos obtidos por esse treinamento podem ser empregados na segmentação dessas imagens.

A Figura 4.12 apresenta o fluxograma da extensão proposta e do método original de Bazen e Gerez [2001]. Note que na modificação apresentada nesta seção, ao contrário do método original, é necessário apenas um treinamento, independentemente da quantidade de sensores considerados.

O capítulo seguinte apresenta alguns experimentos relativos ao método proposto de segmentação.



(a)



(b)

Figura 4.12: Fluxograma do (a) método de Bazen e Gerez [2001] original e da (b) extensão proposta.

Capítulo 5

Experimentos

Este capítulo apresenta os resultados experimentais obtidos com o método de segmentação proposto anteriormente. Nele descreve-se em detalhes os cinco experimentos realizados e as medidas utilizadas na sua avaliação. A Seção 5.1 apresenta os critérios utilizados nas avaliações dos experimentos. As Seções 5.2 e 5.3 mostram os resultados obtidos considerando impressões digitais oriundas de diferentes sensores e impressões digitais roladas, respectivamente. Nas Seções 5.4 e 5.5 avalia-se, respectivamente, o impacto da utilização do método do gradiente (subseção 3.3.1) na *Detecção fina do campo direcional* e a possibilidade da utilização, por outros algoritmos de segmentação, da operação de subtração de *background* definida no Capítulo 4. Na Seção 5.6, por fim, compara-se o método proposto neste trabalho com o apresentado por Yang et al. [2010], que também considera a interoperabilidade de sensor.

5.1 Medidas de avaliação

Com intuito de avaliar de forma objetiva a segmentação realizada pelo método proposto adotou-se a seguinte metodologia. Nos experimentos realizados, a base de referência ou *ground truth* foi obtida a partir da extração manual da região da impressão digital do fundo da imagem, e a exatidão da segmentação foi estabelecida considerando-se as funções *F-score* e de *Jaccard*.

A função *F-score* consiste na média harmônica dos índices de precisão, P , e revocação, R :

$$\text{F-score} = \frac{2 \times P \times R}{P + R} \quad (5.1)$$

em que $P = \frac{tp}{tp+fp}$ e $R = \frac{tp}{tp+fn}$, fp , fn e tp correspondem aos falsos positivos, falsos negativos e verdadeiros positivos obtidos através do *ground truth*.

O resultado desta função varia entre 0 e 1, sendo que o valor 1 representa a segmentação perfeita. Em resumo, a precisão indica a porcentagem de pontos classificados como parte da área da impressão digital, enquanto que a revocação representa a proporção do total de pontos pertencentes à área da impressão que foram classificados corretamente como tal.

A métrica de similaridade *de Jaccard* (ou, ainda, coeficiente de *Jaccard*), por sua vez, é definida de acordo com:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}, \quad (5.2)$$

em que A e B correspondem a dois conjuntos que representam o objeto segmentado na imagem analisada e seu respectivo *ground truth*, e o operador $|X|$ denota a cardinalidade de um conjunto X .

Observe que, nesta função de similaridade, a dualidade (objeto e fundo) da classificação não é considerada, pois somente os pontos da região de interesse são levados em conta no cálculo da similaridade. Dessa maneira, no contexto deste trabalho utilizando essa métrica pode-se ter uma melhor ideia do resultado da segmentação da região das papilas, pois somente a parte da imagem que contém a impressão digital é avaliada e classificadores que não levam em conta essa área não são beneficiados.

Como na função *F-score*, o resultado do *coeficiente de Jaccard* também varia entre 0 e 1 e, do mesmo modo, o valor 1 também representa a segmentação perfeita. Resumidamente, esse coeficiente calcula uma porcentagem para sobreposição da região das papilas entre a imagem analisada e o *ground truth* correspondente.

5.2 Experimento 1

No primeiro experimento, foi considerado um conjunto de imagens típicas de diferentes tipos de sensores, entre eles ópticos, capacitivos e térmicos (Tabela 5.1), obtidas a partir das bases de dados públicas do FVC (*Fingerprint Verification Competition*) [Maio et al., 2002a,b, 2004]. A Figura 5.1 ilustra a diversidade dessas imagens, em particular, das imagens pertencentes ao FVC2004, cujas características são também apresentadas na Tabela 5.1. Observe que, além da diferença de contraste, há também uma diferença de tamanho entre as imagens de bases de dados distintas, uma vez que um mesmo fator de escala foi utilizado para representá-las.

Especificamente, neste trabalho, foram consideradas 778 imagens provenientes das bases de dados FVC2000 (DB3 e DB4), FVC2002 e FVC2004 (DB1, DB2, DB3 e DB4). Os resultados do método proposto foram comparados com os obtidos por aquele apresentado

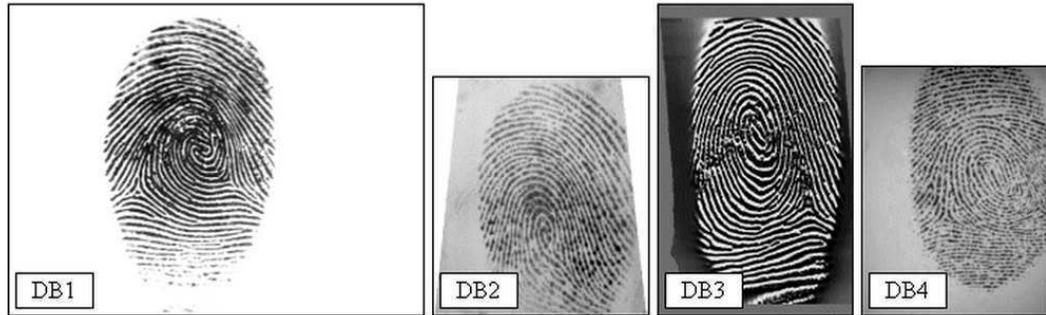


Figura 5.1: Exemplo de imagens retiradas das bases de dados do FVC2004.

Base de dados	Tipo de sensor	Tamanho (pixels)	Resolução
FVC2000 DB3	Sensor óptico “DF-90” - identificador Technology	448 × 478	500 dpi
FVC2000 DB4	Gerador sintético - Sfinge	240 × 320	cerca de 500 dpi
FVC2002 DB1	Sensor óptico “TouchView II” - Identix	388 × 374	500 dpi
FVC2002 DB2	Sensor óptico “FX2000” - Biometrika	296 × 560	569 dpi
FVC2002 DB3	Sensor capacitivo “100 SC” - Precise Biometrics	300 × 300	500 dpi
FVC2002 DB4	Gerador sintético - Sfinge v2.51	288 × 384	cerca de 500 dpi
FVC2004 DB1	Sensor óptico “V300” - Cross-Match	640 × 480	500 dpi
FVC2004 DB2	Sensor óptico “U.are.U 4000” - Digital Persona	328 × 364	500 dpi
FVC2004 DB3	Sensor de varredura térmica “FingerChip FCD4B14CB” - Atmel	300 × 480	512 dpi
FVC2004 DB4	Gerador sintético - Sfinge v3.0	288 × 384	cerca de 500 dpi

Tabela 5.1: Características das imagens pertencentes a base de dados FVC

Base de dados	Bazen e Gerez [2001]	Nosso método
2000-DB3	0.93048	0.96165
2000-DB4	0.97909	0.95933
2002-DB1	0.96828	0.98211
2002-DB2	0.93827	0.95757
2002-DB3	0.97411	0.96044
2002-DB4	0.98215	0.97939
2004-DB1	0.98347	0.98812
2004-DB2	0.92180	0.90778
2004-DB3	0.96060	0.97837
2004-DB4	0.97727	0.96519

Tabela 5.2: Médias da função F-score obtidas para bases de dados FVC

Base de dados	Bazen e Gerez [2001]	Nosso método
2000-DB3	0.79612	0.86103
2000-DB4	0.93559	0.87607
2002-DB1	0.88018	0.92955
2002-DB2	0.85352	0.86857
2002-DB3	0.90495	0.84358
2002-DB4	0.93471	0.92317
2004-DB1	0.87323	0.90743
2004-DB2	0.77547	0.73412
2004-DB3	0.87202	0.92234
2004-DB4	0.92427	0.88469

Tabela 5.3: Médias da função de *Jaccard* obtidas para bases de dados FVC

em Bazen e Gerez [2001] que é um dos mais precisos encontrados na literatura [Maltoni et al., 2009].

Os valores apresentados nas Tabelas 5.2 e 5.3 mostram a equivalência dos resultados das abordagens consideradas. Vale ressaltar a independência do método proposto com relação aos sensores em questão. De fato, o método aqui proposto não inclui qualquer treinamento específico e demonstrou ser robusto o suficiente para lidar com imagens obtidas de uma variedade de sensores, o que implica imagens de diferentes qualidades, resoluções e valores em tons de cinza.

De acordo com as Tabelas 5.2 e 5.3, pode-se observar a equivalência entre os resultados considerando as duas métricas. Isso acontece porque, nesse caso, as duas abordagens segmentam tão bem a região das papilas quanto a do background.

A Figura 5.2 ilustra a diferença entre a segmentação obtida, para um exemplo em

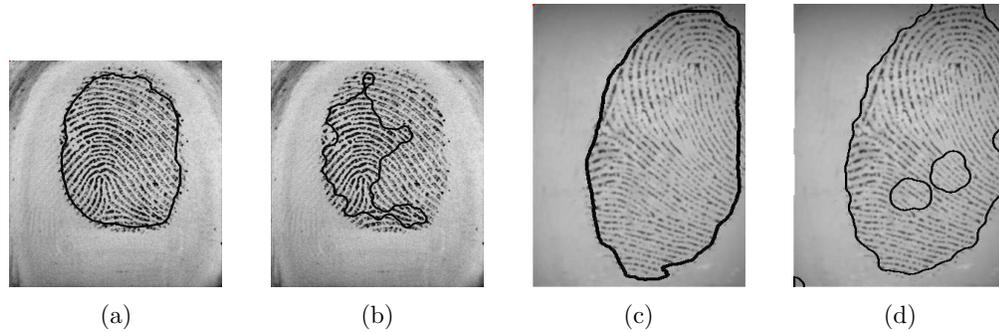


Figura 5.2: Comparação entre métodos: (a) e (c) segmentação obtida pelo método proposto e (b) e (d) segmentação pelo método de Bazen e Gerez [2001].

particular, pelos métodos considerados e a Figura 5.3 ilustra outras segmentações obtidas pelo método proposto considerando as bases de dados do FVC.

Além dos problemas relacionados com a interoperabilidade, mencionados anteriormente, o método de Bazen e Gerez [2001] possui uma outra desvantagem relacionada ao processamento introduzido após a classificação das imagens. Este pós-processamento consiste da aplicação de aberturas e fechamentos necessárias à eliminação de componentes espúrias resultantes. A definição de elemento estruturante para este processo de filtragem não é intuitiva e representa um problema a mais a ser abordado.

5.3 Experimento 2

No segundo experimento, considerou-se um conjunto de 100 imagens de resolução 832×768 pixels disponibilizadas pelo NIST (*National Institute of Standards and Technology*). Neste caso, comparou-se a segmentação obtida levando-se em conta o método de Bazen e Gerez [2001], o sistema biométrico do NIST [Watson et al., 2007] e a abordagem aqui apresentada.

Cada uma das imagens deste experimento corresponde à impressão de um dedo capturada previamente em um cartão de identificação do FBI (*Federal Bureau of Investigation*) e posteriormente digitalizada. Tais imagens apresentam um ruído específico, conforme ilustram as imagens na Figura 5.4. Nessas imagens é possível notar bordas, caracteres provenientes do cartão e anotações manuscritas que devem ser ignorados pelo método de segmentação.

Considerando sua especificidade, o sistema biométrico do NIST, desenvolvido para trabalhar com essa base, realiza um pós-processamento com o objetivo de eliminar esses ruídos. A ideia desse passo adicional é manter o maior componente conexo da imagem. Este pós-processamento foi igualmente aplicado em todos os métodos deste experimento.

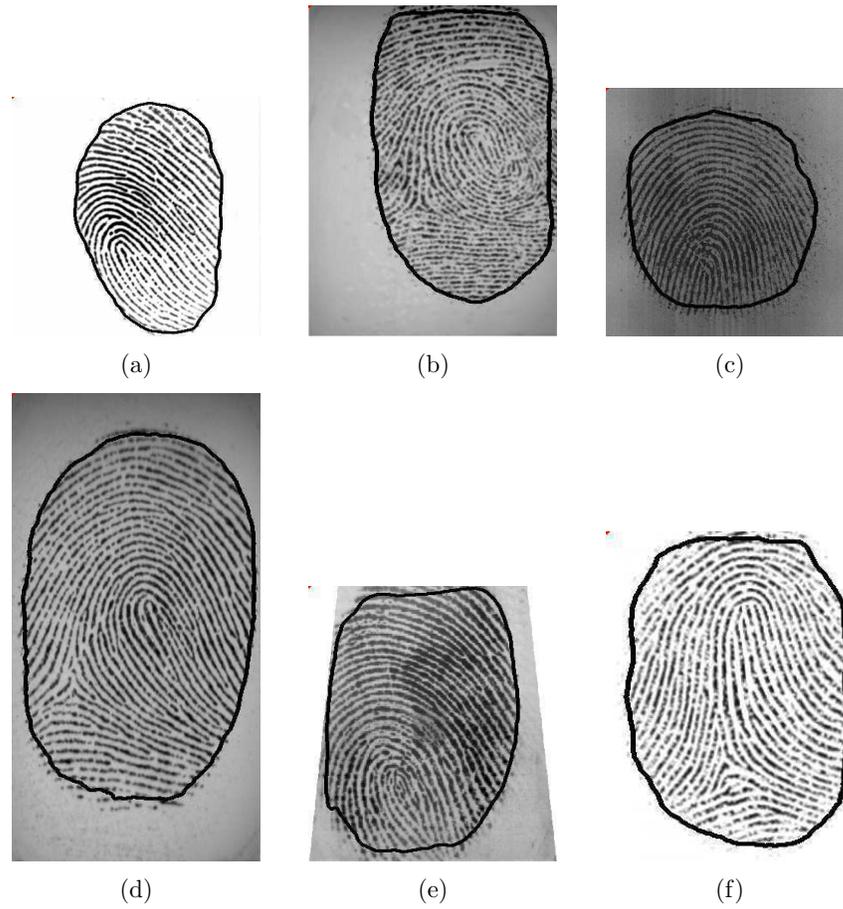


Figura 5.3: Exemplo de segmentação obtida pelo método proposto para imagens das bases de dados do FVC.

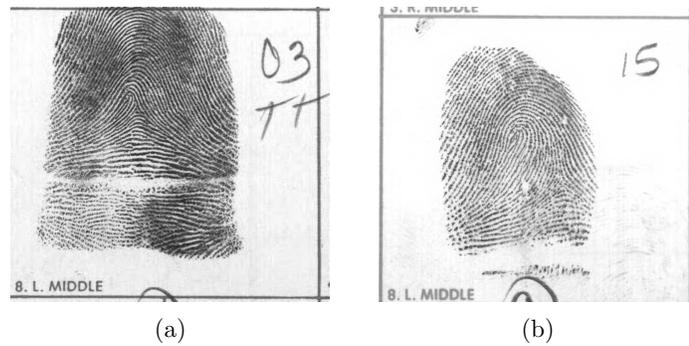


Figura 5.4: Exemplos de imagens da base de dados do NIST.

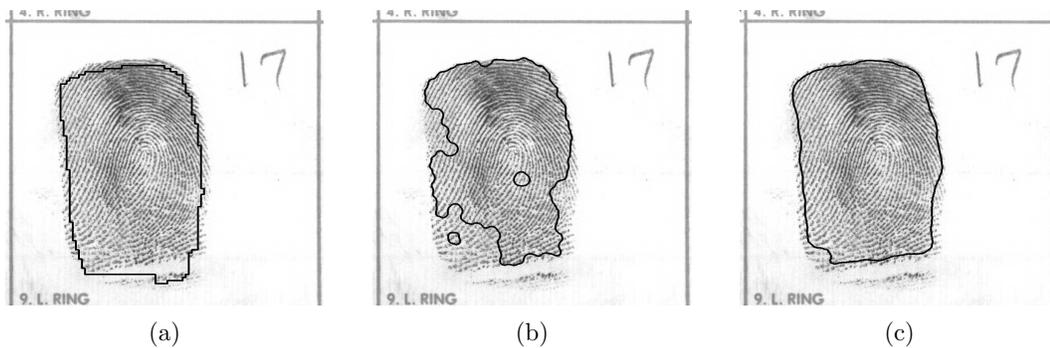


Figura 5.5: Exemplos de segmentação obtida, respectivamente da direita para esquerda, com o sistema biométrico do NIST, o algoritmo de Bazen e Gerez [2001] e o método proposto.

Os resultados são apresentados nas Tabelas 5.4 e 5.5 que indicam uma equivalência de performance do método proposto em relação aos demais. Destaca-se, mais uma vez, a não especificidade da abordagem introduzida com relação à classe de imagens considerada. A Figura 5.5 ilustra alguns exemplos de segmentação.

Uma desvantagem do método de segmentação do NIST é que o mesmo não desconsidera regiões de borrachamentos e, assim, características podem ser eventualmente extraídas destas regiões impróprias, como ilustra a Figura 5.6. Isto acontece porque o método de segmentação proposto por Watson et al. [2007], primeiro, binariza a imagem e, então,

Bazen e Gerez [2001]	NIST [Watson et al., 2007]	Nossa abordagem
0.94193	0.95565	0.95602

Tabela 5.4: Médias da função F-score obtidas para a base de dados do NIST.

Bazen e Gerez [2001]	NIST [Watson et al., 2007]	Nossa abordagem
0.72266	0.789012	0.77903

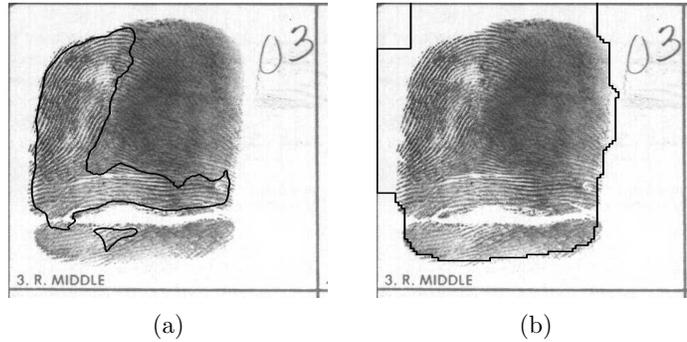
Tabela 5.5: Médias da função de *Jaccard* obtidas para a base de dados do NIST.

Figura 5.6: Exemplos de segmentação com borramento: (a) método proposto e (b) método do NIST.

seleciona as regiões mais escuras para o foreground.

5.4 Experimento 3

No terceiro experimento, avaliou-se a robustez do método utilizado na *Detecção fina do campo direcional*. Para tanto, substituiu-se o método de Oliveira e Leite [2008] pelo método do gradiente (Subseção 3.3.1) e avaliou-se a segmentação resultante.

Nesse caso, levou-se em conta um conjunto de 80 imagens pertencentes à base de dados DB3 do FVC2004. Os resultados obtidos são apresentados nas Tabelas 5.6 e 5.7, que indicam um melhor desempenho do método de detecção de campo direcional introduzido por Oliveira e Leite [2008] em termos da melhoria da segmentação. Uma das razões para esses resultados, é que este método é mais robusto a ruídos presentes na imagem de impressão digital.

Método do gradiente	Oliveira e Leite [2008]
0.96194	0.97837

Tabela 5.6: Médias da função F-score obtidas para base de dados FVC2004 - DB3.

Método do gradiente	Oliveira e Leite [2008]
0.87846	0.92234

Tabela 5.7: Média da função *Jaccard* obtida para base de dados FVC2004 - DB3.

Base de dados	Bazen e Gerez [2001]	Modificação proposta
2004-DB1	0.88731	0.98469
2004-DB2	0.53904	0.95879
2004-DB3	0.97376	0.95263
2004-DB4	0.60655	0.96499

Tabela 5.8: Médias da função F-score obtidas para bases de dados FVC2004.

5.5 Experimento 4

No quarto experimento, avaliou-se a interoperabilidade obtida a partir da modificação do método de Bazen e Gerez [2001] introduzida na Seção 4.5.

Para tanto, considerou-se um conjunto de 308 imagens pertencentes ao FVC2004 [Maio et al., 2004], das quais 20 foram utilizadas no treinamento, sendo 5 de cada uma das bases de dados (DB1, DB2, DB3 e DB4). Primeiramente, o classificador baseado em rede neural foi treinado considerando o conjunto de treinamento e o método de [Bazen e Gerez, 2001] e subsequentemente, outro treinamento foi realizado considerando as imagens obtidas com a subtração de background.

Os valores obtidos neste experimento são apresentados na Tabelas 5.8 e 5.9, que indicam uma melhor performance do classificador treinado ao se considerar a modificação proposta. Tais resultados confirmam a especificidade de sensor (ou base de dados) do espaço de características composto pela média, variância e coerência, discutida por Guo et al. [2009], e mostram o potencial da abordagem proposta em relação ao problema de interoperabilidade de sensor.

Além disso, este experimento mostra o quão significativas são as transformações definidas neste trabalho para segmentação com interoperabilidade, pois avalia o resultado da segmentação utilizando e não utilizando tais transformações.

A Figura 5.7 ilustra os resultados obtidos para uma imagem de cada base de dados considerada.

5.6 Experimento 5

No quinto experimento, foi utilizado o mesmo conjunto de 778 imagens do primeiro experimento. No entanto, neste último, comparou-se os resultados obtidos com o método

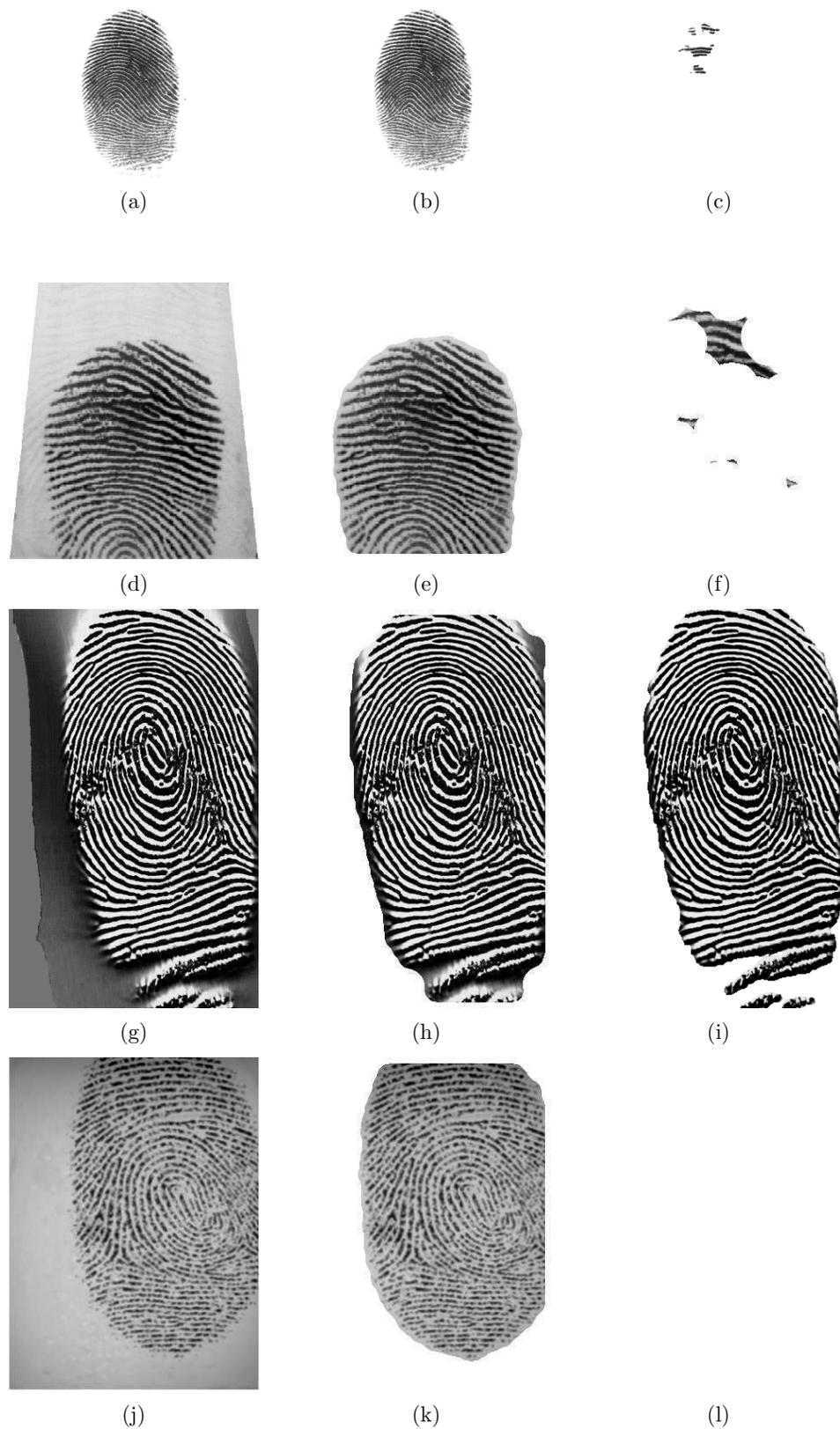


Figura 5.7: Exemplos de segmentação pela modificação proposta: (a), (d), (g) e (j) imagens originais, (b), (e), (h) e (k) modificação proposta e (c), (f), (i) e (l) método de Bazen e Gerez [2001].

Base de dados	Bazen e Gerez [2001]	Modificação proposta
2004-DB1	0.17615	0.88879
2004-DB2	0.00932	0.88509
2004-DB3	0.90972	0.85379
2004-DB4	0.05714	0.89128

Tabela 5.9: Médias da função de *Jaccard* obtidas para bases de dados FVC2004.

Base de dados	Gongping Yang et al. (2010)	Nosso método
2000-DB3	0.95485	0.96165
2000-DB4	0.95763	0.95933
2002-DB1	0.94735	0.98211
2002-DB2	0.94042	0.95757
2002-DB3	0.94781	0.96044
2002-DB4	0.97106	0.97939
2004-DB1	0.95590	0.98812
2004-DB2	0.85707	0.90778
2004-DB3	0.95807	0.97837
2004-DB4	0.96779	0.96519

Tabela 5.10: Médias da função F-score obtidas para bases de dados FVC.

proposto com aqueles alcançados pelo método de Yang et al. [2010].

As Tabelas 5.10 e 5.11 apresentam os valores obtidos e indicam uma melhor performance da abordagem proposta. Como no primeiro experimento, pode-se observar uma equivalência entre os resultados considerando as duas métricas e, do mesmo modo, isso acontece porque, neste caso, as duas abordagens segmentam tão bem a região das papilas quanto a do background.

Como mencionado anteriormente, o trabalho de Yang et al. [2010] é um dos únicos que considera explicitamente a interoperabilidade na segmentação e, assim, estes resultados correspondem a uma das primeiras comparações entre métodos interoperáveis da literatura.

Tal como a abordagem de Bazen e Gerez [2001], o método de Yang et al. [2010] possui uma desvantagem relacionada ao processamento introduzido após o agrupamento dos pontos das imagens. Como no pós-processamento da abordagem em [Bazen e Gerez, 2001] a definição do elemento estruturante para a filtragem no método de Yang et al. [2010] não é intuitiva, o que restringe a interoperabilidade e representa um problema a mais a ser considerado. Vale ressaltar ainda que, neste trabalho, este elemento estruturante foi definido avaliando o resultado da função *F-score* e escolhendo o elemento circular que a maximiza.

Base de dados	Gongping Yang et al. (2010)	Nosso método
2000-DB3	0.83270	0.86103
2000-DB4	0.87415	0.87607
2002-DB1	0.80998	0.92955
2002-DB2	0.82727	0.86857
2002-DB3	0.82521	0.84358
2002-DB4	0.89012	0.92317
2004-DB1	0.65521	0.90743
2004-DB2	0.66811	0.73412
2004-DB3	0.86046	0.92234
2004-DB4	0.89209	0.88469

Tabela 5.11: Médias da função de *Jaccard* obtidas para bases de dados FVC.

Capítulo 6

Conclusão e Trabalhos Futuros

Este trabalho propõe um conjunto de transformações de imagens a serem utilizadas em segmentação e pré-processamento de imagens de impressões digitais, independentemente dos sensores (ou bases de dados) empregados.

Como visto anteriormente, nessas transformações, apenas três parâmetros precisam ser configurados. O primeiro deles, escala, representa o contraste entre diferentes estruturas da imagem, e os dois últimos, número de iterações e tamanho do elemento estruturante planar, se relacionam distância entre cristas.

A partir desse conjunto de transformações, um método de segmentação foi definido e alguns experimentos foram realizados com objetivo de avaliar o resultado da segmentação obtida. Note que, nos experimentos realizados, esses parâmetros foram constantes para todas as bases de dados, o que demonstra, ainda que experimentalmente, a insensibilidade dos mesmos a variações geométricas, topológicas e de escalas de cinza das bases consideradas. Os experimentos com diferentes bases de dados mostram sua equivalência com relação a métodos que representam o *estado-da-arte*, considerando sua exatidão e robustez.

A abordagem proposta, contudo, não sofre de problemas referentes à *interoperabilidade de sensor*, no sentido de que a mesma não requer uma configuração de parâmetros específica, pois características globais da imagem, tais como contraste, são consideradas diretamente pelas transformações introduzidas. De fato, ao contrário de outras abordagens, o método proposto não inclui qualquer treinamento específico e foi desenvolvido para lidar com imagens com diferentes propriedades (resolução, tamanho, qualidade etc).

Os experimentos indicam, ainda, que a abordagem tem potencial para ser aplicada em outros problemas. Algumas das extensões deste trabalho são apresentadas a seguir.

O estudo realizado neste trabalho permite, dentre outras coisas, as seguintes extensões:

- Explorar as propriedades das transformações propostas em outros problemas biométricos com o objetivo de se ampliar a *interoperabilidade de sensor*: as transformações

apresentadas neste trabalho são genéricas o suficiente para serem estendidas a outros problemas relacionados com imagens de impressões digitais, tais como detecção de *poros* e de *pontos singulares*;

- Explorar a representação multi-escala das transformações na definição de um método de avaliação de qualidade de uma imagem de impressão digital: nas transformações definidas, o parâmetro escala se relaciona com o contraste da imagem de tal modo que, à medida que seu valor aumenta, diminui-se também a quantidade de regiões transformadas e apenas regiões com um determinado contraste são consideradas. Pode-se aproveitar essa propriedade para representar regiões com diferentes contrastes e definir para cada uma delas um índice de qualidade.
- Analisar formalmente as propriedades espaço-escala das transformações definidas: neste trabalho foi realizado um estudo experimental e algumas propriedades das transformações foram apenas intuitivamente observadas.
- Utilizar um método para *estimativa de distância entre cristas* na definição dos parâmetros relacionados: na abordagem proposta, apenas três parâmetros precisam ser definidos. No entanto, dois deles, número de iterações e tamanho do elemento estruturante planar, podem ser estimados a partir do tamanho das estruturas da impressão digital, em tempo de execução, utilizando um método que estima a distância entre cristas da impressão.
- Utilizar um algoritmo para escolher a escala adequada considerando o contraste das imagens: se os dois parâmetros relacionados com o tamanho das estruturas da imagem podem ser estimados em tempo de execução, a escala pode ser estimada a partir de observações sobre o conjunto de imagens consideradas. Isso reduziria significativamente o problema da interoperabilidade de sensor.
- Testar o método com uma classe maior de diferentes bases de dados: apesar dos experimentos terem considerado um conjunto relativamente grande de imagens, outros experimentos com outras bases de dados, provenientes de outros tipos de sensores, devem ser realizados.

Referências Bibliográficas

- S. Ando. Consistent gradient operators. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:252–265, 2000. ISSN 0162-8828. doi: <http://doi.ieeecomputersociety.org/10.1109/34.841757>.
- A. M. Bazen e S. H. Gerez. Segmentation of fingerprint images. In *Proc. of the Workshop on Circuits Sys. and Signal Processing*, pages 276–280, 2001.
- A. M. Bazen e S. H. Gerez. Systematic methods for the computation of the directional fields and singular points of fingerprints. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(7):905–919, 2002. 060.02.
- A. M. Bazen, N. J. Bouman e R. N. Veldhuis. A multi-scale approach to directional field estimation. In *Proceedings SAFE & ProRISC 2004 : November 25-26 2004, Veldhoven, the Netherlands*, volume -, pages 215–218, Veldhoven, The Netherlands, 2004. STW/NWO/Dutch Ministry of Economic Affairs. URL <http://doc.utwente.nl/48109/>.
- X. Chen, J. Tian, J. Cheng e X. Yang. Segmentation of fingerprint images using linear classifier. *EURASIP J. Appl. Signal Process.*, 2004:480–494, January 2004. ISSN 1110-8657. doi: <http://dx.doi.org/10.1155/S1110865704309194>. URL <http://dx.doi.org/10.1155/S1110865704309194>.
- L. B. Dorini e N. J. Leite. A scale-space toggle operator for morphological segmentation. In *Proceedings of 8th Int. Symp. on Mathematical Morphology*, volume 1, pages 101–112, 2007.
- L. B. Dorini e N. J. Leite. A multiscale operator for document image binarization. In A. Ranchordas and H. Araújo, editors, *VISAPP (1)*, pages 34–39. INSTICC Press, 2009. ISBN 978-989-8111-69-2.
- R. C. Gonzalez e R. E. Woods. *Digital Image Processing*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2001. ISBN 0201180758.

- C. Gottschlich, P. Mihailescu e A. Munk. Robust orientation field estimation and extrapolation using semilocal line sensors. *Trans. Info. For. Sec.*, 4(4):802–811, 2009. ISSN 1556-6013. doi: <http://dx.doi.org/10.1109/TIFS.2009.2033219>.
- X. Guo, G. Yang e Y. Yin. Sensor interoperability of fingerprint segmentation: An empirical study. In *Information Engineering and Computer Science, 2009. ICIECS 2009. International Conference on*, pages 1–4, 2009. doi: 10.1109/ICIECS.2009.5365419.
- P. Jackway e M. Deriche. Scale-space properties of the multiscale morphological dilation-erosion. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 18(1):38–51, Jan. 1996. ISSN 0162-8828. doi: 10.1109/34.476009.
- A. Jain, Y. Chen e M. Demirkus. Pores and ridges: Fingerprint matching using level 3 features. In *ICPR '06: Proceedings of the 18th International Conference on Pattern Recognition*, pages 477–480, Washington, DC, USA, 2006. IEEE Computer Society. ISBN 0-7695-2521-0. doi: <http://dx.doi.org/10.1109/ICPR.2006.938>.
- A. K. Jain, Y. Chen e M. Demirkus. Pores and ridges: High-resolution fingerprint matching using level 3 features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(1):15–27, 2007. ISSN 0162-8828. doi: <http://doi.ieeecomputersociety.org/10.1109/TPAMI.2007.17>.
- C. Jin e H. Kim. Pixel-level singular point detection from multi-scale gaussian filtered orientation field. *Pattern Recogn.*, 43:3879–3890, November 2010. ISSN 0031-3203.
- M. Kass e A. Witkin. Analyzing oriented patterns. *Comput. Vision Graph. Image Process.*, 37:362–385, March 1987. ISSN 0734-189X. doi: 10.1016/0734-189X(87)90043-0. URL <http://portal.acm.org/citation.cfm?id=27993.27996>.
- J. Koenderink. The structure of images. *Biological Cybernetics*, 50:363–370, 1984. ISSN 0340-1200. URL <http://dx.doi.org/10.1007/BF00336961>. 10.1007/BF00336961.
- T. Lindeberg. Scale-space theory: A basic tool for analysing structures at different scales. *J. of Applied Statistics*, 21(2):224–270, 1994. URL <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.49.4689>.
- D. Maio, D. Maltoni, R. Cappelli, J. Wayman e A. Jain. Fvc2000: fingerprint verification competition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(3):402–412, Mar. 2002a. ISSN 0162-8828. doi: 10.1109/34.990140.
- D. Maio, D. Maltoni, R. Cappelli, J. Wayman e A. Jain. FVC2002: Second Fingerprint Verification Competition. *Pattern Recognition, 2002. Proceedings. 16th*

- International Conference on*, 24:811 - 814 vol.3, 2002b. ISSN 1051-4651. doi: 10.1109/ICPR.2002.1048144.
- D. Maio, D. Maltoni, R. Cappelli, J. L. Wayman e A. K. Jain. Fvc2004: Third fingerprint verification competition. In D. Zhang and A. K. Jain, editors, *Biometric Authentication*, volume 3072 of *Lecture Notes in Computer Science*, pages 1–5. Springer Berlin / Heidelberg, 2004.
- D. Maltoni, D. Maio, A. K. Jain e S. Prabhakar. *Handbook of Fingerprint Recognition*. Springer Publishing Company, Incorporated, 2nd edition, 2009. ISBN 1848822537.
- B. Mehtre. Fingerprint image analysis for automatic identification. *Machine Vision and Applications*, 6:124–139, 1993. ISSN 0932-8092. 10.1007/BF01211936.
- B. Mehtre, N. Murthy, S. Kapoor e B. Chatterjee. Segmentation of fingerprint images using the directional image. *Pattern Recognition*, 20(4):429 – 435, 1987. ISSN 0031-3203. doi: DOI: 10.1016/0031-3203(87)90069-0.
- M. Oliveira e N. Leite. A multiscale directional operator and morphological tools for reconnecting broken ridges in fingerprint images. *Pattern Recognition*, 41(1):367 – 377, 2008. ISSN 0031-3203. doi: DOI: 10.1016/j.patcog.2007.05.019.
- A. Rao e R. Jain. Computerized flow field analysis: Oriented texture fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:693–709, 1992. ISSN 0162-8828. doi: <http://doi.ieeecomputersociety.org/10.1109/34.142908>.
- N. K. Ratha, S. Chen e A. K. Jain. Adaptive flow orientation-based feature extraction in fingerprint images. *Pattern Recognition*, 28(11):1657 – 1672, 1995. ISSN 0031-3203. doi: DOI: 10.1016/0031-3203(95)00039-3.
- A. Ross e A. Jain. Biometric sensor interoperability: A case study in fingerprints. In *in Fingerprints, Appeared in Proc. of International ECCV Workshop on Biometric Authentication*, pages 134–145, 2004.
- A. Ross e R. Nadgir. A calibration model for fingerprint sensor interoperability. In *In SPIE Conference on Biometric Technology for Human Identification III*, 2006.
- J. Serra, L. Vincent e C. D. M. Mathematique. An overview of morphological filtering. In *Circuits, Systems and Signal Processing*, pages 47–108, 1992.
- L. Shen, A. C. Kot e W. M. Koo. Quality measures of fingerprint images. In *Proc. of the 3rd Int. Conf. on Audio- and Video-Based Biometric Person Authentication*, pages 266–271, 2001. ISBN 3-540-42216-1.

- P. Soille. *Morphological Image Analysis*. Springer-Verlag, 1999. ISBN 3-540-65671-5.
- I. Terol-Villalobos, L. Morales-Hernández e G. Herrera-Ruiz. Directional structures detection based on morphological line-segment and orientation functions. In *Discrete Geometry for Computer Imagery*, volume 4992 of *Lecture Notes in Computer Science*, pages 458–469. Springer Berlin / Heidelberg, 2008.
- C. I. Watson, M. D. Garris, E. Tabassi, C. L. Wilson, R. M. McCabe, S. Janet e K. Ko. User's guide to nist biometric image software (nbis). Technical report, NIST Research Report (NISTIR) - 7392, 2007.
- A. Witkin. Scale-space filtering: A new approach to multi-scale description. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '84.*, volume 9, pages 150 – 153, Mar. 1984. doi: 10.1109/ICASSP.1984.1172729.
- G. Yang, G.-T. Zhou, Y. Yin e X. Yang. K-means based fingerprint segmentation with sensor interoperability. *EURASIP Journal on Advances in Signal Processing*, 2010: 1–12, 2010. ISSN 1556-6013. doi: doi:10.1155/2010/729378.
- Q. Zhao, D. Zhang, L. Zhang e N. Luo. High resolution partial fingerprint alignment using pore-valley descriptors. *Pattern Recognition*, 43(3):1050 – 1061, 2010. ISSN 0031-3203. doi: DOI: 10.1016/j.patcog.2009.08.004.
- E. Zhu, J. Yin, C. Hu e G. Zhang. A systematic method for fingerprint ridge orientation estimation and image segmentation. *Pattern Recogn.*, 39(8):1452–1472, 2006. ISSN 0031-3203. doi: <http://dx.doi.org/10.1016/j.patcog.2006.03.001>.
- R. M. McCabe, E. M. Newton, e C. M. Gutierrez. Data format for the interchange of fingerprint, facial and outher biometric information, part one. In *Information*, *American National Standard ANSI/NIST-CSL*, pages 1–150, 2007.