



UNIVERSIDADE FEDERAL DA PARAÍBA  
Programa de Pós-Graduação em Informática

Iron Araújo de Almeida Júnior

# Uma Nova Abordagem de Casamento de Impressões Digitais

João Pessoa

2016



Iron Araújo de Almeida Júnior

# Uma Nova Abordagem de Casamento de Impressões Digitais

Defesa de Mestrado apresentada ao Centro de Informática da Universidade Federal de Paraíba por Iron Araújo de Almeida Júnior, sob a orientação do Prof. Dr. Leonardo Vidal Batista, como parte dos requisitos para obtenção do título de Mestre em Informática.

Orientador: Prof. Dr. Leonardo Vidal Batista

**João Pessoa**  
**2016**

A447u Almeida Júnior, Iron Araújo de.  
Uma nova abordagem de casamento de impressões  
digitais / Iron Araújo de Almeida Júnior.- João Pessoa, 2016.  
113f. : il.  
Orientador: Leonardo Vidal Batista  
Dissertação (Mestrado) - UFPB/CI  
1. Informática. 2. Biometria. 3. Impressão digital.  
4. Impressões digitais - casamento.

UFPB/BC

CDU: 004(043)

Ata da Sessão Pública de Defesa de Dissertação de Mestrado de IRON ARAUJO DE ALMEIDA JUNIOR, candidato ao título de Mestre em Informática, na Área de Sistemas de Computação, realizada em 12 de dezembro de 2016.

1 Aos doze dias do mês de dezembro do ano de dois mil e dezesseis, às dez horas, no Centro de  
2 Informática da Universidade Federal da Paraíba, reuniram-se os membros da Banca  
3 Examinadora constituída para julgar o Trabalho Final do Sr. **IRON ARAUJO DE**  
4 **ALMEIDA JUNIOR**, vinculado a Universidade Federal da Paraíba sob a matrícula nº  
5 2015104014, candidato ao grau de Mestre em Informática, na área de “*Sistemas de*  
6 *Computação*” na linha de pesquisa “*Sinais, Sistemas Digitais e Gráficos*”, do Programa de  
7 Pós-Graduação em Informática. A comissão examinadora foi composta pelos professores:  
8 Leonardo Vidal Batista (PPGI-UFPB), Orientador e Presidente da Banca, Lucídio dos Anjos  
9 Formiga Cabral (PPGI-UFPB), Examinador Interno ao Programa, Jefersson Alex dos Santos,  
10 Examinador Externo à Instituição (UFMG). Dando início aos trabalhos, o Presidente da  
11 Banca, cumprimentou os presentes, comunicou aos mesmos a finalidade da reunião e passou a  
12 palavra ao candidato para que o mesmo fizesse, oralmente, a exposição do trabalho de  
13 dissertação intitulado “*Uma nova abordagem de casamento de impressões digitais*”.  
14 Concluída a exposição, o candidato foi arguido pela Banca Examinadora que emitiu o seguinte  
15 parecer: “*aprovado*”. Do ocorrido, eu, Ruy Alberto Pisani Altafim, Vice-Coordenador do  
16 Programa de Pós-Graduação em Informática (PPGI), lavrei a presente ata que vai assinada  
17 por mim e pelos membros da Banca Examinadora. João Pessoa, 12 de dezembro de 2016.

  
Prof. Dr. Ruy Alberto Pisani Altafim  
Vice-Coordenador do PPGI

**Ruy Alberto Pisani Altafim**  
Vice-Coordenador do Programa de  
Pós-Graduação em Informática  
SIAPE 1971934

Prof. Dr. Leonardo Vidal Batista  
Orientador (PPGI-UFPB)

  
\_\_\_\_\_

Prof. Dr. Lucídio dos Anjos Formiga Cabral  
Examinador Interno ao Programa (PPGI)

  
\_\_\_\_\_

Prof. Dr. Jefersson Alex dos Santos  
Examinador Externo à Instituição (UFMG)

  
\_\_\_\_\_

*Dedico esta conquista ao Sr. Iron Araújo (In memoriam)  
e à Sra. Heloisa Espínola.*

## Agradecimentos

Agradeço à Deus, concebedor de toda sabedoria e graça alcançada até aqui.

À minha mãe Sra. Heloisa Espínola, base da minha estrutura familiar e moral, por me incentivar desde criança e mostrar-me que a educação e o conhecimento possuem valores imensuráveis. Ao meu pai, o Sr. Iron Araújo, que partiu no decorrer deste trabalho, por sempre ter expresso seu amor e me mostrado o bom caminho a ser seguido. À Igreyne Barbosa por ser paciente e sempre me incentivar na busca dos meus sonhos.

Agradeço à Vsoft, na pessoa do Sr. Pedro Alves, por ter me dado a oportunidade de trabalhar em um projeto tão promissor em um ambiente de trabalho agradável. Aos meus companheiros de pesquisa e amigos de verdade Ramon Celestino, Rajiv Albino, Arnaldo Gualberto, Rodrigo (Esquisito) Parente, João Janduy, Igor Andrezza e Erick Cabral e a toda equipe de pesquisa por compartilhar de bons momentos durante esses anos.

Agradeço aos meus orientadores Leonardo Vidal e João Janduy por me guiarem neste caminho tornando-o extremamente interessante. Por compartilhar de seus conhecimentos com paciência e por todas as críticas e esforço coletivo que levaram à produção deste trabalho.

Aos meus companheiros de mestrado e amigos que a vida acadêmica me proporcionou: Matheus Cordeiro, Eduardo Queiroga, Marcílio Lemos, Rubens Godofredo e Daniel Miranda.

Especialmente, a todos os meus mestres, que agregaram valores à minha vida acadêmica e pessoal. E a todos que, de alguma forma, contribuíram para a realização desse objetivo.

O Navio está seguro quando está no porto,  
mas não é para isso que se fazem navios.

*John Shedd.*

## *Resumo*

O reconhecimento de indivíduos por meio de informações biométricas tem sido cada vez mais adotado nas mais diversas aplicações. Uma das razões para essa escolha fundamenta-se no fato de que informações biométricas são mais difíceis de adulterar, compartilhar ou extraviar do que outros métodos de identificação. Várias características de indivíduos, físicas e comportamentais, podem ser utilizadas em um sistema biométrico, como a impressão digital, face, voz e íris. Diversos trabalhos têm explorado esses traços produzindo mecanismos de reconhecimento cada vez mais seguros e precisos, mas nenhum é imune a falhas e ainda há muito o que evoluir. A impressão digital tem se mostrado um traço biométrico bastante eficaz, uma vez que sua alta precisão e seu baixo custo tornam os sistemas mais acessíveis e com resultados satisfatórios. No entanto, o reconhecimento por impressão digital ainda é um problema em aberto, visto que esses algoritmos ainda apresentam erros em sua execução de falsa aceitação e falsa rejeição. Este trabalho apresenta uma nova abordagem de casamento de impressões digitais que é uma etapa importante do processo de identificação. Diversas alternativas de comparação entre as impressões digitais são aplicadas com o objetivo de aumentar a confiabilidade do método proposto. Para ajuste de parâmetros foram utilizados os Algoritmos Bioinspirados: Algoritmos Genéticos e Otimização por Nuvens de Partículas. A validação deste método se fez através da avaliação da Taxa de Erro Igual nos bancos de dados do projeto *Fingerprint Verification Competition* e do projeto BioPass-UFPB, totalizando seis bancos de dados. Resultados apresentaram Taxa de Erro Igual de 0,2%.

**Palavras-chave:** Biometria, Impressão Digital, Casamento de Impressões Digitais

*Abstract*

Biometric recognition has been increasingly adopted in several applications. The main reason for this choice is based on the fact that biometric information is harder to forge, share or misplace in comparison with other authentication alternatives. Many individual traits can be used in a biometric system, such as fingerprint, face, voice, iris and others. Several researches have explored these and other traits producing safer and more accurate recognition methods, but none of them are completely reliable and there are much to improve in this area. This work presents a new approach for fingerprint matching, which is an important step in the identification process. Many comparison alternatives between fingerprints have been used aiming to improve the reliability of the proposed method. This work evaluates two bio-inspired algorithms for parameter optimization, Genetic Algorithms and Particle Swarm Optimization. To validate the proposed method the Equal Error Rate of six databases have been measured, including databases from previous Fingerprint Verification Competitions and a local database named BioPass-UFPA. Results presents 0.2% Equal Error Rate.

**Keywords:** Biometry, Fingerprint, Fingerprint Matching

# Lista de Figuras

2.1	Cristas e Vales de uma impressão digital. Fonte: Maltoni et al. (2009)	9
2.2	Regiões de singularidade (caixas brancas) e pontos centrais (círculos brancos) em impressões digitais. Fonte: Maltoni et al. (2009)	10
2.3	Principais tipos de minúcias. Fonte: Maltoni et al. (2009)	11
2.4	Atributos das minúcias. a) minúcia do tipo terminação: $[x_0, y_0]$ são as coordenadas da minúcia, $\theta$ é o ângulo que a tangente da minúcia forma com o eixo horizontal; b) Minúcia do tipo bifurcação. Fonte: Maltoni et al. (2009)	12
2.5	a) Parte de uma impressão digital, capturada a 1000 DPI, onde os poros ficam bem evidentes; b) Novas cristas; c) vincos e cicatrizes. Fonte: Maltoni et al. (2009)	13
2.6	Alguns sistemas biométricos em um gráfico de segurança versus conveniência. Fonte: Kulp, A. e Braskamp, H. (2012)	13
2.7	Modelo de sistemas biométricos. Fonte: Costa et al. (2006)(Adaptado)	15
2.8	Amostra Pousada. Fonte: Maltoni et al. (2009)	17
2.9	Amostra Rolada. Fonte: Maltoni et al. (2009)	17
2.10	Processo de identificação por impressão digital. Fonte: Adaptado de O'Hara e O'Hara (1956).	18

2.11	As curvas típicas das taxas de erro FAR e FRR, ilustradas uma ao lado da outra, em relação ao limiar T configurado para o sistema. A EER é representada pelo ponto de interseção das curvas. Fonte: Costa et al. (2006). . . . .	21
2.12	Amostras deslocadas, rotacionadas e ruidosas. Fonte: Maltoni et al. (2009).	24
2.13	Amostras diferentes com características semelhantes. Fonte: Maltoni et al. (2009). . . . .	25
2.14	a) Representação de uma minúcia do tipo bifurcação b) Representação de uma minúcia do tipo terminação. . . . .	28
2.15	a) Minúcia do tipo Terminação: (x,y) indica as coordenadas; $\theta$ define a direção média que tangencia os dois vales. b) Minúcia do tipo bifurcação: $\theta$ é definido como a média da direção que tangencia as duas cristas. Fonte: Committees (2005). . . . .	30
2.16	a) e c) Exemplos de pontos em um plano bidimensional contidos em um espaço definidos pela função envoltória convexa. b) Envoltória convexa em um espaço tridimensional. Fonte: (Weisstein, 2016) . . . . .	33
2.17	<i>Leave One Out</i> . Fonte: (Andrezza et al., 2015) . . . . .	34
3.1	Exemplificação dos tipos de cruzamento implementados nos algoritmos genéticos. Fonte: (Silva, 2015) . . . . .	39
3.2	Exemplo da operação de mutação de cromossomos. Fonte: (Silva, 2015)	40
3.3	Partículas em um espaço multidimensional. Fonte: (Coelho Filho et al., 2013) . . . . .	43
3.4	Movimentação de uma partícula. Fonte: (Coelho Filho et al., 2013) . . .	44
3.5	Diagrama de Fluxo do PSO básico. Fonte: (Queiroga, 2016) . . . . .	46
5.1	<i>Workflow</i> do desenvolvimento do método proposto. . . . .	52

5.2	Um exemplo de busca local baseada em vizinhos mais próximos para $K = 2$ . Erros ocorrem devido à minúcias ausentes(b) e informações espúrias(c)	54
5.3	Um exemplo de raio fixo para busca de minúcias na vizinhança. (a) e (b) demonstram o problema decorrente da área de fronteira . . . . .	54
5.4	Informações do Segmento: <i>Reference Angle</i> . . . . .	55
5.5	Informações do Segmento: <i>Neighbor Angle</i> . . . . .	55
5.6	Informações do Segmento: Tamanho do segmento . . . . .	55
5.7	Exemplo de minúcias casadas . . . . .	57
5.8	Exemplo de região de interseção . . . . .	57
5.9	Exemplo de um grafo completo . . . . .	58
5.10	Possível região de interseção entre duas imagens. . . . .	59
5.11	Imagens sobrepostas indicando possível região de interseção. As circunferências azuis indicam minúcias não casadas que pertencem à amostra A, as circunferências vermelhas representam amostras não casadas pertencentes à amostra B e os círculos azuis e vermelhos representam os pares casados das amostras A e B respectivamente. . . . .	61
5.12	Avaliação de minúcias no casamento Global. . . . .	63
A.1	Estrutura de Formação de Templates ISO/IEC 19794-2:2005 de Impressão Digital . . . . .	88
B.1	Classificação dos algoritmos do projeto FVC 2000. . . . .	90
B.2	Classificação dos algoritmos do projeto FVC 2002. . . . .	91
B.3	Classificação dos algoritmos do projeto FVC 2004 banco de dados DB1. . . . .	92
B.4	Classificação dos algoritmos do projeto FVC 2004 banco de dados DB2. . . . .	93
B.5	Classificação dos algoritmos do projeto FVC 2006, classificação do primeiro ao sétimo. . . . .	94

B.6	Classificação dos algoritmos do projeto FVC 2006, classificação do oitavo ao vigésimo quarto colocado. . . . .	94
B.7	Classificação dos algoritmos do projeto FVC 2006, classificação do vigésimo quinto ao trigésimo colocado. . . . .	95

# Lista de Tabelas

2.1	Distribuição, por finalidade, das principais aplicações biométricas. Fonte: Mordini (2005) . . . . .	14
6.1	Casamento local utilizando pesos unitários nas variáveis que compõem a pontuação de similaridade. . . . .	68
6.2	Casamento local e global utilizando pesos unitários. . . . .	68
6.3	Treinamento utilizando Otimização por Nuvens de Partículas . . . . .	70
6.4	Treinamento utilizando Algoritmos Genéticos . . . . .	70
6.5	Método proposto utilizando cálculo de fração sem o cálculo da envoltória convexa. . . . .	71
6.6	Comparação de desempenho no algoritmo proposto utilizando o método da envoltória convexa . . . . .	72
6.7	Desempenho discriminado do algoritmo utilizando as metodologias propostas. . . . .	73
6.8	Comparação de EER's entre algoritmos de casamento de impressões digitais. . . . .	73
6.9	Comparação entre os algoritmos MCC e o método proposto . . . . .	74
6.10	Desempenho do método proposto, de cor verde, em relação aos demais candidatos em todas as versões do projeto FVC. . . . .	76

# Lista de Abreviaturas e Siglas

<b>Símbolo</b>	<b>Descrição</b>
AFIS	<i>Automatic Fingerprint Identification System</i>
CPU	<i>Central Process Unit</i>
EER	<i>Equal Error Rate</i>
FAR	<i>False Accept Rate</i>
FBI	<i>Federal Bureau of Investigation</i>
FRR	<i>False Reject Rate</i>
FVC	<i>Fingerprint Verification Competition</i>
GA	<i>Genetic Algorithm</i>
GB	<i>Gigabyte</i>
GHz	<i>Giga-hertz</i>
MHz	<i>Mega-hertz</i>
OpenCV	<i>Open Computer Vision Library</i>
PSO	<i>Particle Swarm Optimization</i>
SDK	<i>Software Development Kit</i>
SI	<i>Swarm Intelligence</i>
SSD	<i>Solid State Disk</i>

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Objetivos . . . . .	3
1.2	Motivação . . . . .	4
1.3	Estrutura da Dissertação . . . . .	4
<b>2</b>	<b>Fundamentação Teórica</b>	<b>7</b>
2.1	Biometria . . . . .	7
2.2	Impressão Digital . . . . .	9
2.3	Sistemas de Reconhecimento Biométrico . . . . .	12
2.3.1	Sistemas automáticos de Identificação por Impressões Digitais . .	16
2.3.2	Autenticação e Identificação . . . . .	19
2.3.3	Métricas de Erro . . . . .	20
2.4	Casamento de Impressões Digitais . . . . .	22
2.4.1	Casamento de Impressões Digitais baseado em minúcias . . . . .	26
2.5	Padrão ISO de Casamento . . . . .	27
2.6	<i>Fingerprint Verification Competition</i> . . . . .	29
2.7	Agrupamento de Classificadores . . . . .	31
2.8	Envoltória Convexa . . . . .	32
2.9	Validação Cruzada . . . . .	32

<b>3</b>	<b>Computação Bioinspirada</b>	<b>35</b>
3.1	Algoritmos Genéticos . . . . .	36
3.1.1	Conceitos . . . . .	36
3.1.2	Algoritmo . . . . .	39
3.1.3	Parâmetros . . . . .	40
3.2	Otimização por Nuvens de Partículas . . . . .	42
3.2.1	PSO . . . . .	42
<b>4</b>	<b>Trabalhos Relacionados</b>	<b>47</b>
<b>5</b>	<b>Materiais e Métodos</b>	<b>51</b>
5.1	Ambiente de Desenvolvimento . . . . .	51
5.2	Método Proposto . . . . .	52
5.2.1	Casamento Local . . . . .	53
5.2.2	Aplicação do Método Envoltória Convexa . . . . .	58
5.2.3	Casamento Global . . . . .	60
5.2.4	Treinamento . . . . .	63
5.2.5	Bancos de Dados Utilizados . . . . .	65
5.2.6	Validação do Método Proposto . . . . .	66
<b>6</b>	<b>Resultados e Discussão</b>	<b>67</b>
6.1	Desempenho do método proposto utilizando pesos unitários . . . . .	68
6.2	Otimização dos Pesos dos Atributos de Pontuação . . . . .	69
6.2.1	Aplicação do Método Envoltória Convexa . . . . .	71
6.3	Comparação com Outros Trabalhos . . . . .	73
<b>7</b>	<b>Considerações Finais</b>	<b>77</b>
	<b>Referências Bibliográficas</b>	<b>81</b>

Sumário	xvi
<b>A Templates ISO/IEC 19794-2:2005</b>	<b>88</b>
<b>B Classificação dos Participantes nos eventos do projeto FVC.</b>	<b>89</b>

# Capítulo 1

## Introdução

O reconhecimento de indivíduos utilizando suas informações biométricas tem se tornado cada vez mais comum. Os processos desse tipo de identificação foram incluídos no protocolo de acesso a diversas informações e locais privilegiados a fim de aumentar a confiabilidade da permissão daquele que se diz autorizado ao acesso.

O aumento de serviços que necessitam de autenticação gerou uma demanda por novos métodos que possam estabelecer a identidade dos usuários. Métodos tradicionais para estabelecer a identidade de um usuário incluem mecanismos baseados em conhecimento (senhas por exemplo) e mecanismos baseados em *tokens* (cartões inteligentes), que resumidamente representa que o indivíduo é capaz de portar. Porém, os mecanismos de autenticação tradicional podem ser perdidos, esquecidos, roubados ou até mesmo burlados a fim de se passarem pelo seu portador verdadeiro. Neste contexto, a identificação e a autenticação por biometria surgem como alternativas (Ross et al., 2006).

Biometria é definida como a ciência que estabelece a identidade de um indivíduo baseada em suas características físicas, químicas ou comportamentais (Jain et al., 2007). Como esses identificadores biométricos não podem ser facilmente compartilhados ou extraviados, a biometria tem se tornado um componente essencial na construção de

soluções eficazes de identificação pessoal, além do que esses identificadores representam intrinsecamente a identidade corporal dos indivíduos (Maltoni et al., 2009). Do mesmo modo, Jain et al. (2007) afirmam que a biometria estabelece a identidade do indivíduo baseado *no que ele é*, ao invés de *o que ele possui*, como cartões ou dispositivos de validação, ou *do que ele se lembra* como uma senha ou alguma informação de entrada.

O reconhecimento por impressão digital é uma das abordagens biométricas mais utilizadas, uma vez que sua alta precisão e seu baixo custo tornam os sistemas mais acessíveis e com resultados satisfatórios. No entanto, atualmente o reconhecimento por impressão digital ainda é um problema em aberto, visto que erros de falsa aceitação e falsa rejeição ainda são encontrados em seus algoritmos. Esse traço biométrico possui dois tipos de estruturas, as cristas e os vales, dos quais se extraem informações para a realização da comparação entre suas amostras.

Para alcançar uma alta performance no reconhecimento, é essencial a utilização de um algoritmo capaz de extrair informações de uma imagem digital bem como um algoritmo de casamento de impressões digitais (do inglês: *Matching*) robusto, de modo que o conjunto de características encontradas sejam associadas corretamente e haja resistência a ruídos e informações espúrias.

Os sistemas automáticos de identificação de impressões digitais (do inglês: *Automatic Fingerprint Identification System*, AFIS) são os sistemas especialistas quando se trata das digitais e são aceitos e usados em todo o mundo. No início da década de 1960, o FBI começou a desenvolver um AFIS devido à crescente demanda de processos judiciais, o que contribuiu para aumentar o tamanho dos bancos de dados e avançar com essa tecnologia. O sucesso de métodos baseados em impressão digital começou a aumentar rapidamente quando foi criada a tecnologia de *live-scan*, que é uma tecnologia que dispensa o uso de marcas de tinta para aquisição das imagens, no fim da década de 1960 (Raviraj, 2015; Lee et al., 2001).

Outra área da ciência da computação que avança rapidamente é a computação

evolucionária que é relativamente nova. O termo em si foi criado nos anos 1990 e representa um esforço para unir pesquisadores que seguem diferentes linhas simulando vários aspectos da evolução. As técnicas de algoritmos genético, estratégias evolucionárias e programação evolucionária possuem um fundamento em comum: todos envolvem reprodução, variabilidade, competição entre indivíduos dentro de uma população. Com esses aspectos, a evolução é um fim inevitável (Thomas Bäck, 1997).

Ao longo dos anos, os algoritmos evolucionários têm provado ser uma ferramenta significativa na busca de soluções em problemas complexos, como: caixeiro viajante, escalonamento de eventos, detecção de objetos, entre outros (Dawid, 2011; Sun et al., 2004; McLachlan et al., 2002). Na seleção de características, a aplicação desses algoritmos é capaz de melhorar a taxa de precisão dos classificadores; melhorar a velocidade e a eficiência dos preditores através da otimização de características; e fornecer uma melhor compreensão do processo subjacente que gerou os dados (Luque-Baena et al., 2013).

Neste trabalho utiliza-se algoritmos evolucionários com a finalidade de minimizar a taxa de erro no algoritmo proposto através da otimização de seus parâmetros de medição de similaridade.

## 1.1 Objetivos

O objetivo do presente trabalho é criar um algoritmo baseado em uma nova abordagem de casamento de impressão digital. Essa abordagem unirá as vantagens coletadas em vários trabalhos presentes da literatura, além de pesquisas desenvolvidas independentemente, e utilizará algoritmos evolucionários com o objetivo de otimizar os parâmetros do algoritmo e diminuir a taxa de erro nos bancos de dados utilizados.

Os objetivos específicos são:

- Desenvolver um algoritmo de casamento de impressões digitais.
- Comparar a eficácia da aplicação de diferentes algoritmos bioinspirados utilizados

na otimização de parâmetros para avaliação de similaridade entre amostras.

- Analisar a aplicação de técnicas diferentes às aquelas presentes na literatura.

## 1.2 Motivação

A identificação utilizando impressões digitais é uma área da biometria bastante explorada e uma das mais antigas dessa ciência. Apesar disso, as pesquisas em âmbito nacional ainda são bastante escassas. Como pode-se observar em BioLab (2006), que é uma competição de referência mundial e verifica o desempenho de algoritmos de identificação através de impressão digital, na modalidade FM-ISO, pode-se constatar a ausência de algoritmos produzidos no Brasil. Isso implica em perda de espaço e de oportunidades neste país, quando se trata de mercado comercial público e privado, onde é priorizado produtos pesquisados e desenvolvidos dentro do país (Lei de Licitações, 8.666 de 1993; Instrução Normativa 04, 2010). Dessa forma, um algoritmo igualmente eficiente e de boa qualidade, como os produzidos em outros países, terá prioridade em licitações e no uso nacional em geral, como por exemplo: Sistemas biométricos utilizados em votações eleitorais, fornecidos pelos Tribunais Eleitorais, que são sistemas que atuam em grande escala de uso.

Além disso, a literatura ainda apresenta trabalhos que são deficientes quando se trata de capturas de impressões digitais em casos especiais, como capturas de impressões digitais de baixa qualidade ou capturas de impressões digitais parciais.

## 1.3 Estrutura da Dissertação

Na duas seções seguintes, Capítulo 2 e Capítulo 3, será apresentada a fundamentação teórica necessária para o entendimento deste trabalho. Em seguida, no Capítulo 4, tratar-se-á de informações presentes na literatura, tanto na parte de casamento quanto em avaliação de métodos de treinamento de algoritmos. No Capítulo 5, o algoritmo

proposto será apresentado juntamente com os métodos de avaliação que foram utilizados e os materiais de pesquisa. No Capítulo 6 serão apresentados os resultados e a discussão. E no Capítulo 7 serão feitas as considerações finais deste trabalho.



## Capítulo 2

# Fundamentação Teórica

Neste Capítulo são apresentados os principais conceitos teóricos empregados na realização desse trabalho. Primeiramente, são apresentados conceitos acerca de biometria e algumas de suas vertentes. Logo após, definições a respeito de impressões digitais e sistemas biométricos e outras informações necessárias ao entendimento deste trabalho.

### 2.1 Biometria

O termo Biometria refere-se ao uso de características físicas ou comportamentais, tais como face, íris, impressão digital, voz, forma de digitar e forma de escrever para identificar pessoas automaticamente. Segundo Clarke (1994), qualquer característica humana pode ser usada como característica biométrica desde que satisfaça alguns requisitos básicos:

- **Universalidade:** deve ser uma característica presente em toda a população. Na prática, existem pessoas que não têm impressões digitais ou mãos, por exemplo.
- **Unicidade:** uma característica biométrica deve ser única para cada indivíduo, ou seja, a possibilidade de pessoas distintas possuírem características idênticas, deve ser nula ou desprezível. Assim, a altura de uma pessoa não é uma boa

característica para autenticação, já que várias pessoas podem ter a mesma altura. Na prática, as características biométricas podem apresentar maior ou menor grau de unicidade, mas nenhuma delas pode ser considerada absolutamente única para cada indivíduo. Segundo Costa et al. (2006), a quantidade de variação devida à genética e ao ambiente muda de biometria para biometria. Cada pessoa é única, se analisada com suficiente detalhe. É próximo do impossível que duas pessoas diferentes tenham a mesma representação biométrica em qualquer sistema razoável. Contudo, ao lidar com tecnologias práticas de autenticação, encontramos limites na resolução das imagens extraídas, na capacidade de armazenamento e na habilidade de comparação entre dados extraídos. Na prática, isto extermina a noção de unicidade absoluta para todas as características biométricas.

- Permanência: a característica deve ser imutável. Na prática, existem algumas alterações causadas pelo envelhecimento, pela mudança das condições de saúde ou mesmo emocionais das pessoas e por mudanças nas condições do ambiente de coleta.
- Coleta: a característica tem que ser passível de mensuração por meio de um dispositivo. Na prática, todas as características biométricas utilizadas comercialmente atendem a este requisito.
- Aceitação: a coleta da característica deve ser tolerada pelo indivíduo em questão. Na prática, existem preocupações com higiene, privacidade e questões culturais que diminuem aceitação da coleta.

É importante salientar que, na prática, nenhuma característica atende com perfeição a todos os requisitos de uma característica biométrica ideal.

Para o uso de uma característica para fim de reconhecimento biométrico, alguns critérios devem ser atendidos: o processo de identificação através da característica deve possuir um desempenho aceitável, no que diz respeito à precisão e velocidade; é

preciso saber até que ponto as pessoas aceitam determinado identificador biométrico em suas vidas, disponibilizando determinada característica para sua identificação; é preciso observar também o quão fácil é enganar o sistema com técnicas fraudulentas (Jain et al., 2004)

## 2.2 Impressão Digital

Impressão digital é uma reprodução da aparência exterior da epiderme do dedo. A estrutura mais característica de uma impressão digital é o padrão intercalado entre as cristas (do inglês: *Ridges*) e os vales (do inglês: *Valleys*). Em uma captura de impressão digital, as cristas são as áreas escuras enquanto que os vales são as áreas claras, Figura 2.1. As cristas possuem diferentes espessuras, variando de 100  $\mu\text{m}$ , para as mais finas até, 300  $\mu\text{m}$  para as mais largas. A maior parte das lesões encontradas no dedo, como queimaduras superficiais, abrasões ou cortes, não afetam a estrutura de saliência das cristas, e o padrão original é duplicado por toda a nova pele que cresce (Maltoni et al., 2009).

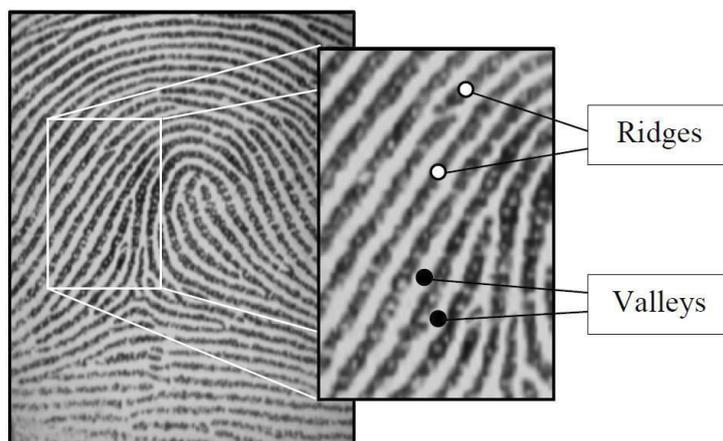


Figura 2.1: Cristas e Vales de uma impressão digital. Fonte: Maltoni et al. (2009)

Os detalhes das cristas são geralmente descritos em uma ordem hierárquica de três

níveis diferentes: Nível 1 (Padrão Global de disposição das cristas), Nível 2 (Minúcias), e Nível 3 (Poros, formato das arestas locais das cristas).

No nível global (Nível 1), as cristas geralmente se encontram paralelamente umas às outras, porém existem uma ou mais regiões onde elas assumem formas distintas (grande curvatura, maior quantidade de terminações, etc.). Essas regiões, chamadas de singularidades ou regiões singulares, são classificadas em três topologias: presilha (do inglês: *Loop*), delta e verticilo (do inglês: *Whorl*) (Figura 2.2). Existem situações em que uma região verticilo não é apresentada pois podem ser descritas como duas presilhas, uma de frente para a outra (Maltoni et al., 2009).

Algoritmos de comparação de impressões digitais podem pré-alinhar as imagens baseados em um ponto central, chamado de núcleo (do inglês: *Core*), que pode ser definido como o centro da presilha mais ao norte da imagem. Existem, no entanto, imagens que não possuem singularidade do tipo presilha ou verticilo, nesse caso o núcleo é geralmente associado ao ponto máximo de curvatura das cristas. Contudo, devido à grande variação dos padrões das impressões digitais, muitas vezes é uma tarefa difícil encontrar um núcleo confiável em todas as amostras.

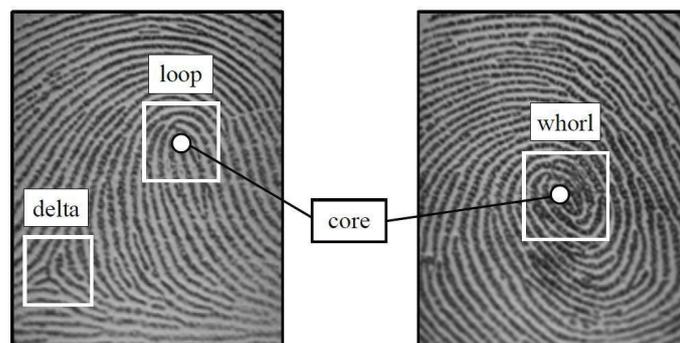


Figura 2.2: Regiões de singularidade (caixas brancas) e pontos centrais (círculos brancos) em impressões digitais. Fonte: Maltoni et al. (2009)

No Nível 2, em um escopo menor, outras características podem ser encontradas, as

minúcias. Minúcia, que significa pequeno detalhe, no contexto das impressões, digitais refere-se às diferentes maneiras que uma crista pode ser descontinuada. Uma crista pode terminar subitamente, são as terminações (do inglês: *Ridge Ending*), ou pode sofrer uma bifurcação e se dividir em duas cristas, são as bifurcações (do inglês: *Ridge Bifurcation*). As minúcias são as características mais utilizadas nos algoritmos de comparação de impressões digitais. Em uma impressão digital completa, onde a captura é feita por meio de rolagem do dedo, o número de minúcias encontradas pode ser maior que 100, no entanto um pequeno número de minúcias corretamente correspondidas é suficiente para afirmar que as duas amostras são originadas do mesmo dedo. Existem diversos tipos diferentes de minúcias (Figura 2.3), no entanto apenas alguns tipos são utilizados (Maltoni, 2009).

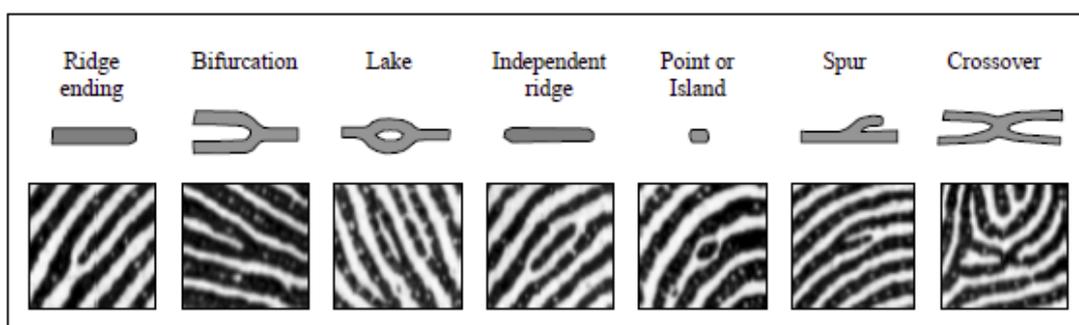


Figura 2.3: Principais tipos de minúcias. Fonte: Maltoni et al. (2009)

O modelo utilizado pelo FBI (polícia federal dos Estados Unidos), baseado nas coordenadas das minúcias, utiliza apenas terminações e bifurcações de cristas. Cada minúcia é classificada pelas suas coordenadas nos eixos X e Y, e pelo ângulo da minúcia que tangencia as cristas (Figura 2.4). Foi adicionado o ângulo da minúcia como um atributo aos algoritmos de comparação de impressões digitais devido à falta de precisão das minúcias capturadas, onde em algumas amostras, a variação da pressão do dedo sobre o sensor poderia dificultar a identificação de bifurcações e terminações, confundindo-as.

Em um nível muito mais local está o Nível 3, onde detalhes mais finos podem

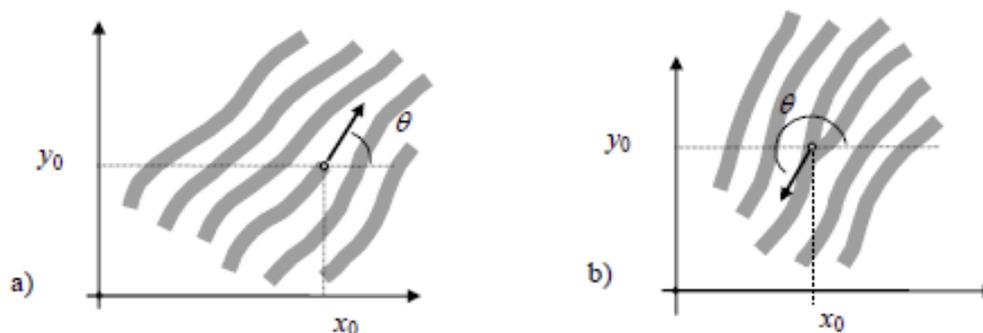


Figura 2.4: Atributos das minúcias. a) minúcia do tipo terminação:  $[x_0, y_0]$  são as coordenadas da minúcia,  $\theta$  é o ângulo que a tangente da minúcia forma com o eixo horizontal; b) Minúcia do tipo bifurcação. Fonte: Maltoni et al. (2009)

ser extraídos de uma impressão digital. Nesse nível estão incluídos todos os atributos dimensionais das cristas, como largura, formato, contorno das bordas, poros, cristas novas, vincos e cicatrizes. Cada crista é dotada de poros em toda sua extensão. O tamanho dos poros pode variar entre  $60 \mu\text{m}$  e  $250 \mu\text{m}$ . Estudos revelam que o número de poros por centímetro de crista varia 9 a 18 e estima-se que entre 20 e 40 poros são suficientes para determinar a identidade de um indivíduo. As características extraídas no Nível 3, apesar de serem extremamente distintivas e importantes para examinadores de impressões digitais, são pouco utilizadas nos algoritmos de comparação automática, uma vez que para obter uma detecção confiável são necessários sensores de alta resolução, que são mais caros, e impressões digitais de boa qualidade. Na Figura 2.5 é possível observar os detalhes de uma impressão digital no Nível 3.

### 2.3 Sistemas de Reconhecimento Biométrico

Jain et al. (2004) afirmam que um sistema biométrico é essencialmente um sistema de reconhecimento de padrões que opera através da aquisição de dados biométricos de um indivíduo, extraíndo um conjunto de características a partir dos dados adquiridos e comparando-o com o conjunto de modelo no banco de dados. Uma vez que os identifi-

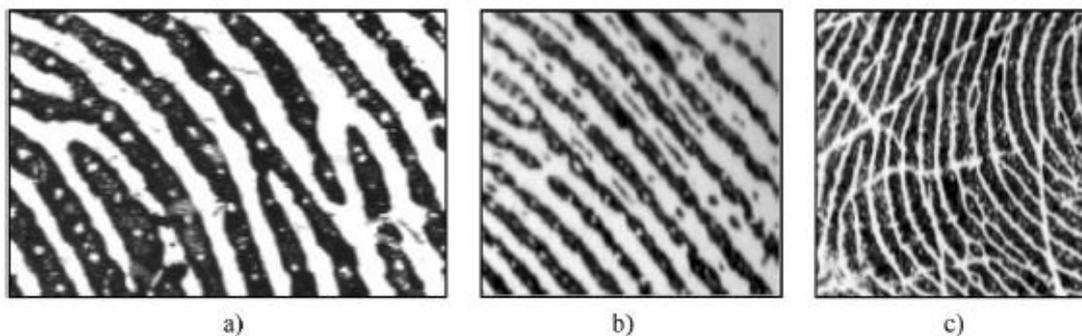


Figura 2.5: a) Parte de uma impressão digital, capturada a 1000 DPI, onde os poros ficam bem evidentes; b) Novas cristas; c) vincos e cicatrizes. Fonte: Maltoni et al. (2009)

cadres biométricos não podem ser facilmente extraviados, forjados, ou compartilhados, métodos de identificação biométricos são considerados mais confiáveis do que métodos baseados em *tokens* (como *smartcards*) ou senhas (Maltoni et al., 2009). Dessa forma, sistemas de reconhecimento biométrico estão sendo cada vez mais utilizados em várias aplicações governamentais, civis e forenses.



Figura 2.6: Alguns sistemas biométricos em um gráfico de segurança versus conveniência. Fonte: Kulp, A. e Braskamp, H. (2012)

Os sistemas biométricos têm diversas utilidades, entre elas, o aumento na segurança, prevenção do roubo de identidade, identificação e verificação de indivíduos e proteção de ativos nacionais e organizacionais, além da autenticação para utilização de aplicativos

(Rahman et al., 2008). Na Figura 2.6, Kulp, A. e Braskamp, H. (2012) trazem uma relação entre a confiabilidade dos métodos biométricos, quando se trata de segurança, e a conveniência de seu uso, que pode ser traduzida como aceitabilidade dos usuários.

Tabela 2.1: Distribuição, por finalidade, das principais aplicações biométricas. Fonte: Mordini (2005)

Finalidade	Utilização
Identificação Criminal	28%
Controle de acesso e atendimento	22%
Identificação Civil	21%
Segurança de redes e de computadores	19%
Autenticação em pontos de vendas, ATM's e varejos	4%
Autenticação telefônica e comércio eletrônico	3%
Vigilância e filtragem	3%

De uma maneira geral, os sistemas biométricos podem ser divididos em aplicações de nicho Governamental, Comercial e Forense que podem ser estendidas para a classificação por finalidade. Mordini (2005) dividiu essas finalidades em sete grupos e mensurou a utilização de cada uma delas (Tabela 2.1).

Um sistema biométrico pode ser encarado como um sistema de reconhecimento de padrões de propósito específico (Bolle et al., 2002). Independente da finalidade, seu modelo conceitual leva em consideração os dados e processos básicos comuns a qualquer sistema biométrico. Em uma primeira utilização, o usuário é previamente registrado e seu perfil biométrico fica armazenado. Na utilização posterior do sistema, o processo de aquisição obtém os dados biométricos apresentados e características particulares dos dados são extraídas para comparação com o perfil armazenado. O processo de comparação decide se os dados apresentados são suficientemente similares ao perfil registrado (Costa et al., 2006).

Costa et al. (2006) define essas etapas:



Figura 2.7: Modelo de sistemas biométricos. Fonte: Costa et al. (2006)(Adaptado)

- **Aquisição e exemplar:** O processo de aquisição ou apresentação é o processo de obtenção dos dados da característica biométrica oferecida. Normalmente, a dificuldade deste processo é balancear adequadamente a qualidade da amostra sem causar excesso de inconveniência para o usuário. Neste módulo é geralmente embutido um controle da qualidade da amostra adquirida (viabilidade de processamento). O exemplar, ou amostra, é o resultado do processo de aquisição.
- **Extração e atributos:** O processo de extração produz uma representação computacional do exemplar obtido, que chamaremos de atributos, ou características extraídas. A extração de características é a redução de um conjunto de medidas formado por uma grande quantidade de dados que contém uma pequena quantidade de informação útil para um conjunto que contém menos dados, mas praticamente a mesma quantidade de informação (Patrick, 1972).
- **Registro e perfil:** O processo de registro (do inglês: *Enrollment*), obtém previamente os dados biométricos do usuário para cadastramento no sistema. O perfil biométrico obtido, ou *template*, é armazenado para uma comparação posterior. O processo de registro é necessário para o estabelecimento do perfil para posterior comparação.
- **Comparação, limiarização e decisão:** O processo de comparação verifica qual é o grau de similaridade entre as características extraídas da amostra do usuário e o perfil armazenado previamente. Este processo fornece um escore representativo

da similaridade entre os dois conjuntos de dados. Caso a similaridade seja superior a certo limite previamente determinado, conhecido como limiar (do inglês: *Threshold*), a decisão é aceitar o usuário, ou seja, uma autenticação válida. Caso a similaridade seja inferior ao limiar, a decisão é não aceitar o usuário, e então temos um usuário não autenticado.

### 2.3.1 Sistemas automáticos de Identificação por Impressões Digitais

Historicamente, a captura das impressões digitais era realizada no “método de tinta”, nele os dedos eram melados em tinta preta e pressionados ou rolados sobre um cartão branco; depois o cartão era digitalizado por *scanners* de propósito geral. Esse tipo de processamento era conhecido como captura *offline* de impressões digitais.

A coleta de impressões digitais deixadas em cenas de crime é um exemplo de uso da captura *offline*. Hoje em dia, a maioria dos sistemas automáticos de identificação por impressão digital (AFIS) de propósito civil e criminal utilizam *live-scans*, e obtém imagens digitais a partir da aplicação direta de sensores eletrônicos sobre superfícies que contenham impressões digitais, ou da aplicação de dedos sobre os sensores. Nenhuma tinta é necessária para esse método e tudo o que se deve fazer é colocar o dedo no sensor.

As metodologias usadas pelos AFIS têm se beneficiado com a obtenção das digitais por *live-scanners*, essas inovações foram importantes para a aceitabilidade e conveniência de aplicações comerciais e civís, pois nelas não se espera que seja necessário aplicar tinta nos dedos todas as vezes que o usuário desejar utilizar algum serviço.

Os sensores lêem padrões das cristas encontradas na superfície do dedo, convertendo a leitura analógica para uma imagem de impressão digital. Existem diferentes tecnologias de sensores como: óticos, ultrassom, entre outros. Alguns sensores suportam a captura de impressões digitais do tipo roladas (Figura 2.9), nas quais o usuário rola o dedo horizontalmente com o objetivo de capturar a maior quantidade de informação

possível. Esse tipo de captura obtém mais informações do que a captura pousada (Figura 2.8), aquela em que o usuário apenas pousa o dedo sobre o sensor (Maltoni et al., 2009).



Figura 2.8: Amostra Pousada. Fonte: Maltoni et al. (2009)



Figura 2.9: Amostra Rolada. Fonte: Maltoni et al. (2009)

Os sistemas de autenticação por impressão digital funcionam conforme é descrito na Figura 2.10. Inicialmente uma amostra é capturada a partir de um sensor específico,

essa amostra poderá ser usada para cadastro da impressão digital ou poderá ser usada para buscar por uma digital previamente cadastrada. O cadastro se dá através do armazenamento do *template* que é obtido após o processo de extração de características da imagem. Se o usuário estiver cadastrado no banco de dados, através de comparações, este poderá ser identificado, a depender do propósito da aplicação.

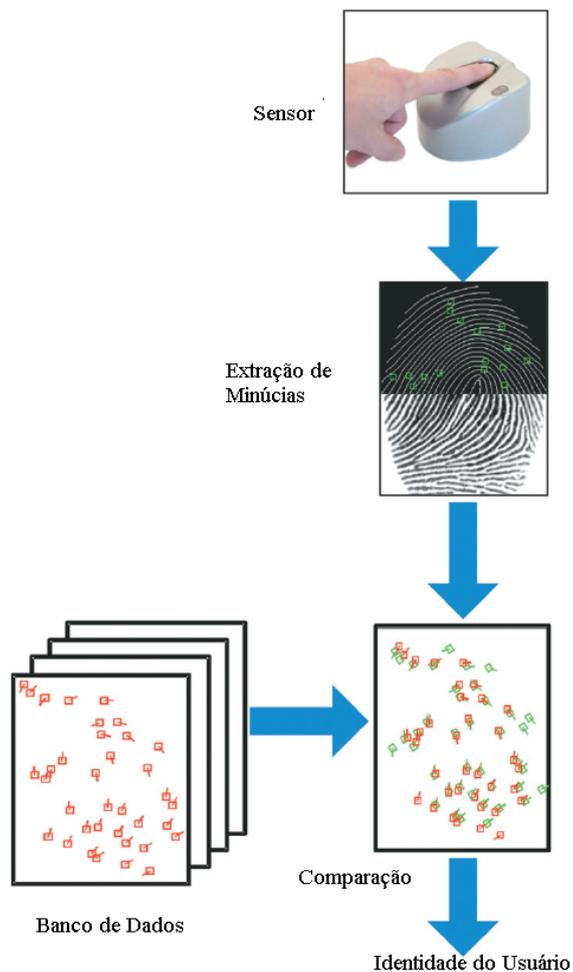


Figura 2.10: Processo de identificação por impressão digital. Fonte: Adaptado de O'Hara e O'Hara (1956).

### 2.3.2 Autenticação e Identificação

Os sistemas biométricos são usados para a autenticação de pessoas. Nestes sistemas existem três modos de autenticação: a verificação, a classificação e a identificação (Bolle, 2004).

Na verificação, a característica biométrica é apresentada pelo usuário juntamente com a identidade alegada, usualmente por meio da digitação de um código de identificação. Esta abordagem de autenticação é dita uma busca um-para-um (1:1), ou busca fechada, em um banco de dados de perfis biométricos e é tipicamente utilizada em aplicações de reconhecimento positivo, cujo objetivo é prevenir que múltiplas pessoas usem a mesma identidade (Jain et al., 2004). O princípio da verificação está fundamentado na resposta à questão: "O usuário é quem diz que é?".

Na classificação, por sua vez, dada uma característica biométrica pelo usuário, o sistema busca todos os registros do banco de dados e retorna uma lista de registros com características suficientemente similares à característica biométrica apresentada. O resultado pode ser refinado posteriormente por comparação adicional, biometria adicional, ou intervenção humana (Costa et al., 2006). Neste método utiliza-se comparação do tipo um-para-muitos (1-N) são realizadas. Por conta disso, geralmente a precisão diminui com o aumento da base de dados. De acordo com Jain et al. (2004) a classificação é um componente crítico em aplicações de reconhecimento de negativos, onde o sistema verifica se uma pessoa é quem ela nega ser. O propósito de reconhecimento negativo é prevenir que uma pessoa tenha múltiplas identidades. Basicamente, a classificação corresponde à questão: "Quem é o usuário?".

Por fim, ainda há outro conceito utilizado em sistemas biométricos: a identificação. Desta vez, o objetivo do sistema é determinar se uma dada amostra está presente na base de dados utilizada e, em caso afirmativo, definir a quem pertence através da classificação. A identificação, além de não ser comumente implementada em sistemas biométricos, é também considerada um problema mais difícil, pois o sistema agora

deve verificar se a amostra em questão é diferente de todas as demais, já que ela não foi cadastrada previamente. Na identificação, o sistema deve responder à questão: "O usuário está cadastrado?".

### 2.3.3 Métricas de Erro

O desempenho da verificação de sistemas biométricos geralmente é medido utilizando métricas como taxa de *Falsa Aceitação* (do inglês: *False Acceptance Rate*, FAR), taxa de *Falsa Rejeição* (do inglês: *False Rejection Rate*, FRR) e taxa de *Erro Igual* (do inglês: *Equal Error Rate*, EER). A falsa aceitação ocorre quando um usuário não registrado obtém êxito ao tentar acessar um sistema protegido, enquanto a falsa rejeição ocorre quando um usuário registrado não consegue obter o acesso legítimo por não ter sido encontrado no banco de dados.

De acordo com Ross et al. (2006), as taxas FAR e FRR podem ser calculadas pelas Equações 2.1 e 2.2:

$$FAR(l) = \int_l^{\infty} p(s|impostor)ds \quad (2.1)$$

$$FRR(l) = \int_{\infty}^l p(s|genuine)ds \quad (2.2)$$

Onde  $p(s|impostor)$  e  $p(s|genuine)$  representam distribuições de probabilidade de um escore  $s$  sobre condições de genuíno e impostor, respectivamente, e  $l$  é um limiar que define se um indivíduo é genuíno ou impostor.

Além disso, o sistema pode operar nas faixas de conveniência ou de segurança, conforme calibração do limiar. Todas essas medidas são inversamente proporcionais em relação à qualidade do desempenho do sistema biométrico.

É importante destacar que o ajuste do limiar permite a calibragem do sistema, mas também implica consequências opostas. Sistemas onde brechas na segurança devem

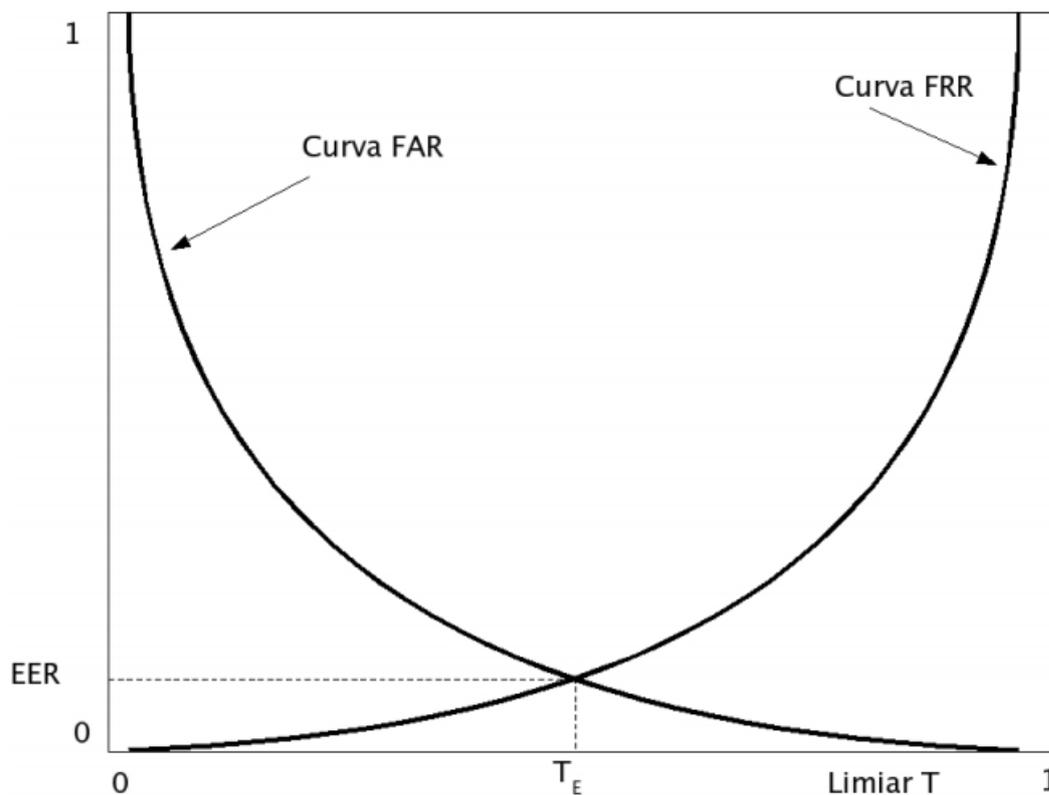


Figura 2.11: As curvas típicas das taxas de erro FAR e FRR, ilustradas uma ao lado da outra, em relação ao limiar  $T$  configurado para o sistema. A EER é representada pelo ponto de interseção das curvas. Fonte: Costa et al. (2006).

ser evitadas ao máximo (e.g., sistemas bancários), ajustam o limiar para que a taxa FAR seja mínima, operando na faixa de segurança. No entanto, isso pode resultar em problemas de conveniência, pois usuários genuínos poderão ter seu acesso negado até que o sistema tenha a total certeza sobre o perfil do usuário. Por outro lado, sistemas que buscam minimizar a FRR no sentido de melhorar a conveniência de utilização por parte do usuário, poderão ter problemas de segurança, autorizando o acesso de usuários não legítimos (Silva, 2015).

A EER é geralmente utilizada para comparar o desempenho de diferentes sistemas biométricos (Zhang, 2004) e é calculada pelo ponto de interseção entre as curvas FAR e FRR (Figura 2.11).

## 2.4 Casamento de Impressões Digitais

Um algoritmo de casamento de impressões digitais (do inglês: *Fingerprint Matching*) compara duas amostras e indica um grau de similaridade, geralmente um valor numérico entre 0 e 1 ou um valor booleano indicando se as impressões digitais pertencem ao mesmo dedo ou não. Apenas alguns algoritmos de casamento atuam diretamente nas imagens em escala de cinza geradas pelos sensores. A maioria deles necessitam de uma representação intermediária decorrente da extração de características de uma dada digital.

A comparação de imagens de impressões digitais geralmente é difícil, dada a grande variabilidade nas diferentes amostras da mesma impressão digital. Segundo Maltoni et al. (2009), os principais fatores que causam essas variações podem ser divididos em:

- **Deslocamento:** Um mesmo dedo pode ser posicionado em diferentes posições ao tocar o sensor durante a captura das digitais, resultando na translação de toda a imagem.
- **Rotação:** Um mesmo dedo pode ser rotacionado em diferentes ângulos ao ser posicionado sobre um sensor para a captura das amostras. Essas rotações involuntárias podem ter uma variação de até 40°.
- **Captura Parcial:** Uma translação ou rotação na captura geralmente resultam em perda de parte do dedo que pode sair da área do sensor, resultando em uma área pequena de extração do dedo do usuário. Esse problema é mais sério quando as áreas dos sensores são menores.
- **Distorções Não Lineares:** O sensoriamento mapeia a forma tridimensional de um dedo sobre a superfície de bidimensional do sensor. Esse mapeamento resulta em uma distorção não linear em diferentes aquisições das amostras de um dedo devido a elasticidade da pele. Geralmente os algoritmos de casamento descartam

algumas características desse mapeamento e consideram a imagem não distorcida da digital.

- **Pressão e Condicionamento da Pele:** A estrutura de uma crista pode ser precisamente capturada se as cristas da parte do dedo que toca o sensor tiverem um contato uniforme. Entretanto, a pressão do dedo, a aridez da pele, doenças de pele, suor, sujeira, gordura e a umidade do ar alteram esse estado, resultando em um contato não uniforme. As consequências são ruídos e diferentes capturas de amostras do mesmo dedo, que piora com a magnitude desses fatores
- **Ruído:** Resíduos deixados sobre o vidro do sensor proveniente de capturas anteriores.
- **Erros na Etapa de Extração de Características:** Os algoritmos de extração de características são imperfeitos e geralmente apresentam falhas. Esses erros podem ser provenientes durante qualquer etapa da extração. Em imagens de baixa qualidade, o processo de extração de minúcias pode resultar em vários erros a ainda não detectar todas as minúcias existentes (Maltoni et al., 2009).

A Figura 2.12 mostra exemplos de digitais que apresentam as características apresentadas.

Por outro lado, amostras de pessoas diferentes podem parecer bastante similares, especialmente se tratando da estrutura global da digital (posições das singularidades, orientação das cristas). Entretanto a possibilidade de um grande número de minúcias de impressões de dois dedos diferentes “casarem” é extremamente pequena. No casamento de impressões digitais busca-se o melhor alinhamento entre as amostras e também igualar minúcias de amostras diferentes mesmo que elas não estejam perfeitamente alinhadas (Figura 2.13).

Existem muitos algoritmos de casamento de digitais propostos na literatura. A maior parte desses algoritmos não têm dificuldade ao comparar imagens de boa qualidade.



Figura 2.12: Amostras deslocadas, rotacionadas e ruidosas. Fonte: Maltoni et al. (2009).

Entretanto, o casamento de impressões digitais ainda é um desafio quando aplicado a imagens de baixa qualidade ou a imagens parciais de digitais. No caso de AFIS' assistidos por pessoas, um algoritmo de checagem de qualidade pode ser usado para que apenas imagens de boa qualidade sejam usadas e armazenadas nos bancos de dados. Porém, a ação humana não é viável em sistemas de reconhecimento biométricos autônomos, os quais são cada vez mais crescentes no mercado.

Uma análise de erros de Falsas Rejeições produzidas por vários algoritmos de casamento que participaram do FVC2000 mostraram que a maior parte dos erros ocorreram em 20% das imagens de baixa qualidade. Em outras palavras, 20% do banco de dados é responsável por 80% das Falsas Rejeições. Avanços no estado da arte no reconhecimento de digitais foram observados nas edições do FVC (Maltoni et al., 2009).

Ainda é necessário um desenvolvimento contínuo para sistemas mais robustos que



Figura 2.13: Amostras diferentes com características semelhantes. Fonte: Maltoni et al. (2009).

sejam capazes de processar e comparar imagens de baixa qualidade, em particular para aplicações que lidam com esse tipo de imagem.

Segundo Maltoni et al. (2009) os algoritmos de impressões digitais pode ser classificados em três famílias:

- Casamento baseado em Correlação: Duas imagens de impressões digitais são superpostas e a correlação entre os *pixels* é calculada em diferentes alinhamentos (rotação, translação...).
- Casamento baseado em Minúcias: Esse é o método mais popular e mais usado, e se tornou a base para comparação de papiloscopista. As minúcias são extraídas

das duas digitais e armazenadas como um conjunto de pontos num plano bidimensional. Esse método consiste em encontrar um alinhamento entre os templates obtidos e resulta no maior conjunto de minúcias pareadas obtidas.

- Casamento não baseado em Minúcias: A extração de minúcias é extremamente difícil em imagens de baixa qualidade. Enquanto algumas outras características de padrões das cristas podem ser extraídas mais seguramente do que as minúcias (como formato das cristas ou informações de textura da imagem), a sua captura é, menos distinguível assim como a persistência em outras amostras é menor. A abordagem desse tipo de casamento compara duas amostras em termos de características extraídas de padrões encontrados nas cristas.

#### 2.4.1 Casamento de Impressões Digitais baseado em minúcias

Chikkerur (2005) ainda divide os algoritmos de casamento de impressões digitais baseado em minúcias em algumas categorias.

- O casamento local, é uma abordagem de comparação que aborda estruturas locais, na vizinhança das minúcias. Cada vizinhança é associada a propriedades estruturais que são invariantes à rotação e translação. Entretanto, a vizinhança local não captura informações acerca do relacionamento global entre estruturas em um escopo global, o que pode gerar falsos casamentos. A confiabilidade deste método está em combinar todos os casamentos obtidos a partir destas estruturas que são encontradas em escopos menores.
- O casamento global é uma abordagem de comparação que tenta, simultaneamente, alinhar todos os pontos de interesse em um único passo. Em outras palavras, a função de transformação assumida deve alinhar todos os pontos da impressão digital. Entretanto, a transformada é apenas uma aproximação pois sabe-se que capturas

diferentes da mesma impressão digital são distorcidas pelas transformações não lineares.

Esse alinhamento ainda pode ser dividido em duas categorias.

- O Alinhamento Implícito, que é usado na maioria dos métodos de casamento (Chikkerur, 2005), realiza o alinhamento e a busca pela correspondência dos pontos como um processo único.
- O Alinhamento Explícito, utiliza uma transformada obtida através de outro processo de alinhamento e em seguida utiliza a informação obtida para parear as demais minúcias.

Por convenção, adotaremos a representação das minúcias como na Figura 2.14. Onde o centro da circunferência representa a localização da minúcia e a direção do segmento de reta identifica a direção da minúcia. Quando a extensão da circunferência possuir apenas um segmento, indicará que a minúcia é do tipo terminação e quando possuir dois segmentos indicará que a minúcia é do tipo bifurcação e a direção da minúcia será a reta que passa entre os dois segmentos.

## 2.5 Padrão ISO de Casamento

O padrão ISO/IEC 19794-2:2005 especifica o formato dos dados para as representações de impressões digitais baseado em minúcias. Define uma forma genérica para o seu armazenamento de forma que possa conter informações de uma ou mais amostras de impressões digitais. Esse padrão define informações importantes como: o tipo das minúcias, a posição e a orientação, e ainda especifica a forma que os dados serão armazenados.

Segundo Maltoni et al. (2009), esse padrão ajuda a promover a interoperabilidade de sistemas, visto que a extração de características de uma imagem digital provenientes

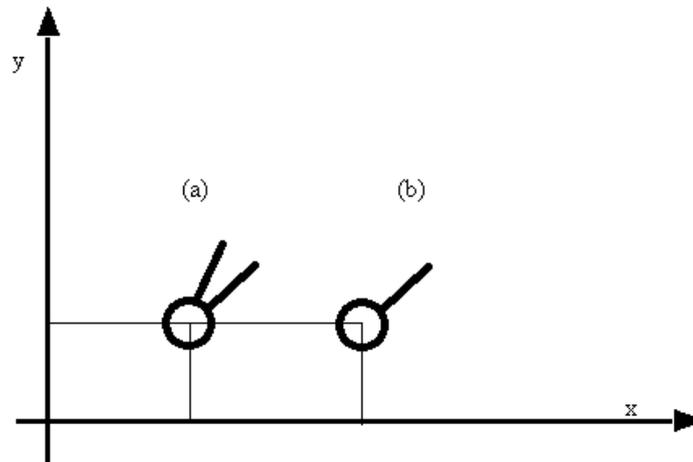


Figura 2.14: a) Representação de uma minúcia do tipo bifurcação b) Representação de uma minúcia do tipo terminação.

de um sistema que é capaz de gerar *templates* nesse formato permite que *templates* que seguem o padrão ISO possam ser usados por algoritmos de casamento de impressão digital diferentes. Esse tipo de padrão costuma ser usado em bases de dados compartilhadas de instituições governamentais, acadêmicas ou em competições. Nesse tipo de base de dados, independente do algoritmo de extração de características, os algoritmos de casamentos são capazes de realizar comparações entre os diferentes tipos de *templates*. O padrão ISO utiliza apenas informações da abordagem local baseada em minúcias.

O formato de gravação baseado em minúcias define os elementos fundamentais usados para a representação de impressões digitais baseada em minúcias e outros dados opcionais como: contagem de cristas, informações acerca dos núcleos e das demais singularidades. O Apêndice A sumariza as informações que estão contidas em um template no padrão ISO através de uma tabela.

No cabeçalho desse arquivo ainda existem informações gerais a respeito da imagem, como: o tamanho da imagem, a quantidade de amostras presentes. Onde para cada amostra existe informações sobre suas minúcias e informações adicionais.

Para cada minúcia, o espaço que é destinado a seu armazenamento contém:

- Tipo de Minúcia: O tipo da minúcia pode ser do tipo terminação e bifurcação. Além dos dois tipos de minúcias conhecidas, um tipo definido como "outro" identifica um tipo de minúcia que não pôde ser identificado com precisão, podendo pertencer a ambos os grupos ou ser tratada como uma minúcia desconhecida.
- Posição: A posição X e Y é expressa em *pixels* de acordo com as coordenadas do sistema (Figura 2.15).
- Orientação: A orientação  $\theta$  da minúcia é gravada em unidades de *bytes*, conforme segue:  $1.40625(360/256)$  graus.
- Qualidade: A qualidade da minúcia indica a certeza da precisão de sua informação. Essa, por sua vez, é expressa entre 1 (qualidade mínima) e 100 (qualidade máxima) ou 0 se não há informações definidas de qualidade.

Outras informações sobre a contagem de cristas e núcleos das digitais também podem ser gravadas a fim de serem usados para aumentar a confiabilidade dos algoritmos de casamento de impressões digitais.

## 2.6 *Fingerprint Verification Competition*

O FVC (*Fingerprint Verification Competition*) é uma competição internacional que avalia algoritmos utilizados no reconhecimento de impressão digital. Essa competição possui alguns bancos de dados desconhecidos pelos participantes e visa avaliar o desempenho entre os algoritmos a fim de ordená-los por desempenho. Assim é possível avaliar os algoritmos com melhores desempenho e aqueles que estão no estado da arte.

O FVC já contou com quatro versões nos anos: 2000, 2002, 2004 e 2006. Nestas versões os competidores submetiam seus algoritmos a avaliação em datas definidas.

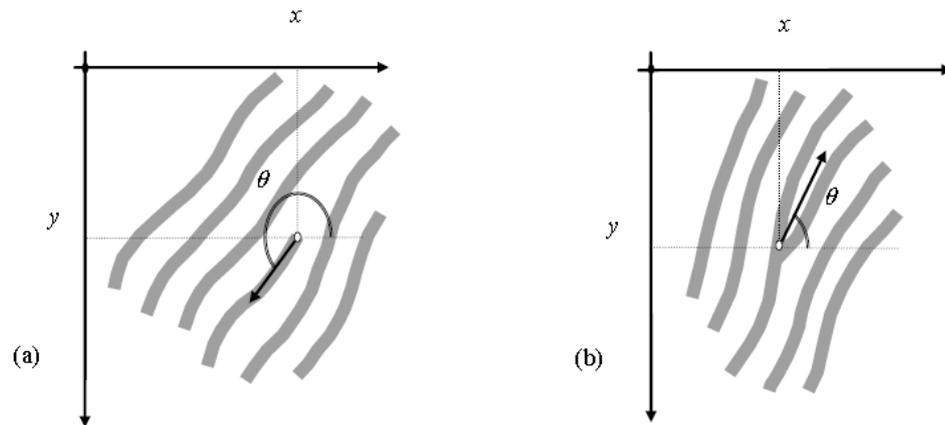


Figura 2.15: a) Minúcia do tipo Terminação:  $(x,y)$  indica as coordenadas;  $\theta$  define a direção média que tangencia os dois vales. b) Minúcia do tipo bifurcação:  $\theta$  é definido como a média da direção que tangencia as duas cristas. Fonte: Committees (2005).

Atualmente a versão FVC-onGoing está em andamento, nesta versão os competidores não precisam esperar uma data pré-determinada para que seus algoritmos sejam avaliados. A avaliação acontece de modo "ongoing", ou seja, os algoritmos podem ser submetidos a qualquer tempo que serão avaliados em uma mesma base de dados e o *ranking* de classificação altera-se a cada novo algoritmo, que tem desempenho superior aos anteriormente classificados, é submetido à competição.

Os participantes da competição devem disponibilizar arquivos separadamente para as etapas de extração de características e de casamento; a avaliação é conduzida a partir desses arquivos sobre um conjunto de banco de dados criados a partir do mesmo conjunto de sensores que também criaram o conjunto de testes, que está disponível aos participantes.

O FVC é organizado pelo laboratório de sistemas biométricos da Universidade de Bolonha (*Biometric System Laboratory, University of Bologna*); laboratório de reconhecimento de padrões e processamento de imagens, da Universidade do Estado de Michigan (*Pattern Recognition and Image Processing Laboratory - Michigan State Uni-*

versity); centro de testes biométricos da Universidade do Estado de San Jose (*Biometric Test Center - San Jose University*) e pelo grupo de reconhecimento biométrico da Universidade Autônoma de Madrid (*Biometric Recognition Group - Universidad Autónoma de Madrid*). A primeira, segunda, terceira e quarta edição do FVC receberam atenção das comunidades acadêmicas e industriais relacionadas com biometria (BioLab, 2006).

## 2.7 Agrupamento de Classificadores

Os métodos de agrupamentos de classificadores (do inglês: *Ensemble Methods*) são algoritmos de aprendizagem de máquina que constroem um conjunto de classificadores e então classificam uma nova instância baseado em uma votação ponderada desses preditores.

Essas decisões dos classificadores são combinadas de tal modo que seja possível classificar novos exemplos. A principal descoberta realizada nos estudos acerca dos *ensembles* é que eles são muito mais preciso do que os classificadores que o formam (Dietterich, 2000).

A condição necessária e suficiente para que os agrupamentos sejam mais eficientes que os classificadores individualmente é se os classificadores forem precisos e diversificados (Hansen e Salamon, 1990). Um classificador é dito preciso se sua taxa de acerto é melhor do que uma classificação aleatória, ou seja, maior que 50% de acertos. E dois classificadores são ditos diversos se eles cometem erros diferentes em instâncias a serem classificadas. Para verificar o porquê a diversidade e a precisão são bons, Dietterich (2000) dá o seguinte exemplo: dado três classificadores que formam um *ensemble*  $h_1$ ,  $h_2$ ,  $h_3$  e considere um novo caso  $x$ . Se os três classificadores são idênticos, e não diversos, então quando  $h_1(x)$  errar, então  $h_2(x)$  e  $h_3(x)$  também errarão. Entretanto se os erros causados pelos classificadores não forem correlacionados então se  $h_1(x)$  errar, eventualmente  $h_2(x)$  e  $h_3(x)$  acertarão e o voto da maioria classificará a nova instância

corretamente.

Neste trabalho utiliza-se o agrupamento de classificadores para algoritmos de casamento de impressões digitais, com a finalidade de aumentar a confiabilidade do método proposto.

## 2.8 Envoltória Convexa

O problema envoltória convexa (do inglês: *Convex Hull*) foi um dos primeiros problemas do campo de geometria computacional estudado, do ponto de vista de complexidade computacional.

A função envoltória convexa de um conjunto finito de pontos  $s$  em um plano é o menor polígono convexo que contém todos os pontos de  $s$ . Os vértices desse polígono devem ser pontos que estão contidos no conjunto  $s$ . Para calcular a função envoltória convexa é necessário encontrar os pontos que serão parte dos vertices limitadores do polígono. (Kirkpatrick e Seidel, 1986)

Segundo o'Rourke (1998) é possível implementar a função envoltória convexa com complexidade  $O(n^2)$  em um espaço tridimensional. Essa função é capaz de trabalhar eficientemente em um espaço de 2 a 8 dimensões (Barber et al., 1996).

A Figura 2.16 exemplifica a aplicação do método em conjuntos de pontos em espaços dimensionais variados. A Figura 2.16(a) mostra um conjunto de pontos após a aplicação da envoltória convexa criando um polígono que contém todos os pontos. O mesmo pode ser feitos em estruturas tridimensionais como mostra a Figura 2.16(b).

## 2.9 Validação Cruzada

Para validar um classificador deve-se testar a sua eficácia. A validação cruzada é uma técnica muito utilizada para isso quando há um número reduzido de amostras para treinamento e testes. Para tal, o banco de dados utilizado é dividido em vários grupos

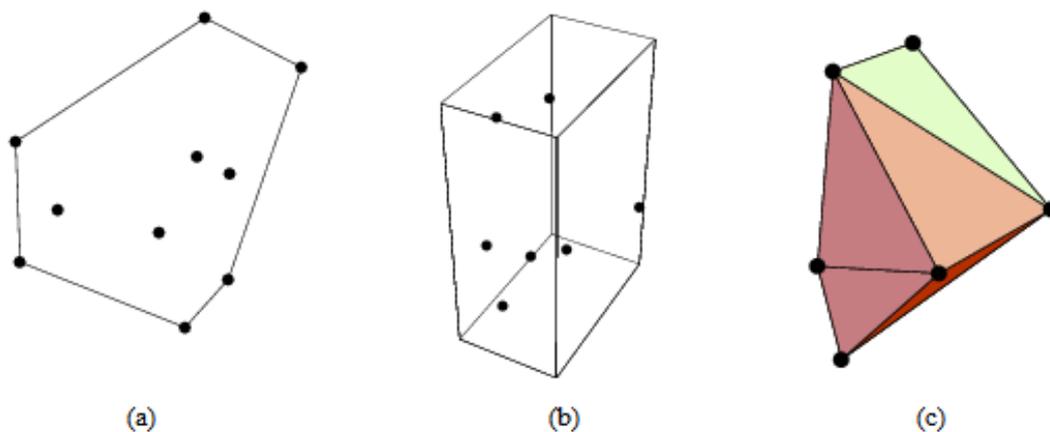


Figura 2.16: a) e c) Exemplos de pontos em um plano bidimensional contidos em um espaço definidos pela função envoltória convexa. b) Envoltória convexa em um espaço tridimensional. Fonte: (Weisstein, 2016)

disjuntos, um para treinamento e os outros para testes de classificação. Várias rodadas de classificação são realizadas com diferentes divisões dos grupos, a fim de se obter um resultado médio.

Um caso especial da Validação Cruzada é a técnica do deixe-um-de-fora (do inglês: *Leave-One-Out*). Como o próprio nome sugere, essa técnica deixa uma amostra de cada classe fora do conjunto de treinamento. Esta será a amostra de teste ou de classificação, e o treinamento é realizado com as demais. O processo é repetido  $n$  vezes, onde  $n$  é a quantidade de amostras no banco de dados, até que todas as amostras do banco tenham ficado de fora do treinamento. A média dos resultados de classificação destas repetições será o resultado final.

A Figura 2.17 ilustra a abordagem do deixe-um-de-fora, onde a cada rodada, um elemento (em cinza) é classificado, e os demais elementos (em branco) são utilizados no treinamento. A cada rodada modifica-se o conjunto de treinamento e o elemento a ser classificado, até que todos os elementos tenham ficado uma vez de fora (Andrezza et al., 2015).

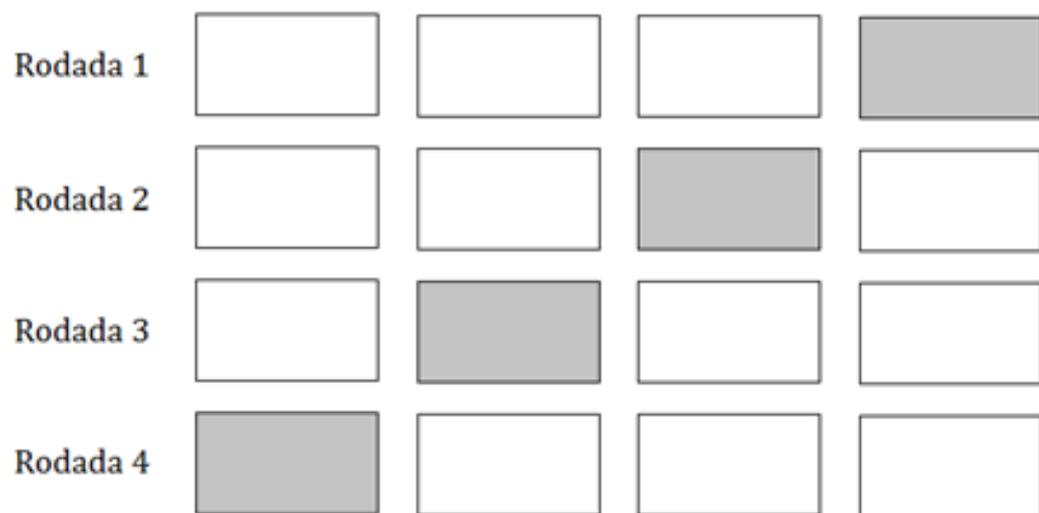


Figura 2.17: *Leave One Out*. Fonte: (Andrezza et al., 2015)

## Capítulo 3

# Computação Bioinspirada

A computação bioinspirada é a linha de pesquisa que emprega metáforas e modelos de sistemas biológicos no projeto de ferramentas computacionais de solução de problemas complexos (De Castro e Von Zuben, 2005; Paton, 1994). Essa classe de meta-heurísticas tem recebido bastante atenção nos últimos tempos e é constituída por ferramentas de computação bioinspirada ou computação inspirada na biologia.

Muitos desses processos biológicos têm um comportamento otimizado e uma capacidade de processamento de informação muito maior e melhor do que o melhor computador já construído até hoje. A prova disso é todo o processo de evolução que as espécies tiveram ao longo de milênios, assim como a capacidade do cérebro humano de guardar informações em forma de imagens, cheiros, sabores, sons; bem como realizar cálculos lógico e matemáticos (França, 2005).

Um dos primeiros algoritmos bioinspirados originou-se da tentativa de entender o funcionamento do sistema nervoso com o trabalho de McCulloch e Pitts (1943), que resultou no primeiro modelo matemático do funcionamento do neurônio biológico. Mais tarde, já com Rosenblatt (1957), foi desenvolvida a ideia que criou a primeira rede neural capaz de aprender a classificar padrões linearmente separáveis.

A seguir serão apresentados dois algoritmos bioinspirados, o primeiro é baseado na

teoria da evolução das espécies de Darwin (1872), os algoritmos genéticos (do inglês: *Genetic Algorithms*, GA) e o segundo é um algoritmo de inteligência de enxames que é baseado em comportamento de populações que é observado em toda a natureza, a otimização por nuvens de partículas (do inglês: *Particle Swarm Optimization*, PSO).

### 3.1 Algoritmos Genéticos

Os algoritmos genéticos foram inicialmente propostos por Holland (1962) como uma ferramenta de busca de soluções em problemas de otimização. Eles são baseados no princípio da seleção natural descrito por Darwin (1872). Darwin decreve, em seu livro, que a aptidão é, basicamente, a capacidade de sobreviver. E isso faz muito sentido quando aplicado a problemas onde a aptidão de um indivíduo é dada pelo valor de uma função de adaptação que será otimizada.

Na computação, os algoritmos genéticos são simples de se descrever e programar, embora seu comportamento em tempo de execução possa ser complicado de ser visualizado. Por causa disso, alguns trabalhos têm sido realizados para explicar como o GA funciona e como se aplica (Holland, 1975; Golberg, 1989; Whitley, 1993). Em termos gerais, acredita-se que os algoritmos genéticos funcionam pela descoberta, ênfase e recomendação de boas soluções que tendem a recombinar-se e formar soluções de melhor aptidão.

Em seguida, apresentar-se-ão os principais conceitos acerca dos algoritmos genéticos.

#### 3.1.1 Conceitos

Tipicamente, um algoritmo genético trabalha numa população de indivíduos. Cada indivíduo é representado por um cromossomo formado por um conjunto de genes representando os parâmetros a serem otimizados. Certas operações são realizadas com o objetivo de produzir novas gerações de indivíduos baseado nas suas capacidades de

gerar melhores resultados, são elas:

- **Cruzamento:** é a operação chave para gerar novos indivíduos na população. Inspirado pelo exemplo da natureza, o cruzamento é projetado para juntar material genético de cromossomos com um alto valor de aptidão a fim de produzir indivíduos mais aptos. Usualmente, os algoritmos genéticos implementam três tipos de cruzamento (Figura 3.1), tais quais:
  - **Um ponto:** no cruzamento de um ponto, os cromossomos pais são divididos aleatoriamente em duas partes pelo mesmo gene e, então, os cromossomos filhos são gerados pela junção das partes dos pais.
  - **Dois pontos:** semelhantemente ao cruzamento de um ponto, nesse caso, cada cromossomo pai é dividido em três partes pela escolha aleatória de dois genes. O processo de geração do cromossomo filho também é semelhante ao cruzamento de um ponto.
  - **Uniforme:** o cruzamento uniforme realiza aleatoriamente o cruzamento dos cromossomos pais na criação dos filhos. Primeiramente, os cromossomos pais são divididos em duas ou mais partes aleatórias na primeira geração, cujas divisões são repetidas nas próximas gerações. Então, os filhos são criados pela união das partes genéticas dos pais.
- **Seleção:** o operador de seleção destina-se a implementar a ideia de sobrevivência do mais apto. Esse operador basicamente determina qual dos cromossomos na população atual tem permissão para herdar o seu material genético na próxima geração. Há 4 tipos usuais de métodos de seleção:
  - **Roleta:** nesse tipo de seleção, os cromossomos são selecionados de acordo com sua aptidão. Quanto mais apto o cromossomo, maior é a chance de ele ser selecionado para o cruzamento em um sorteio.

- Classificação: na seleção por classificação, os cromossomos são, primeiramente, classificados por seu valor de aptidão. Assim, o pior cromossomo recebe o valor de aptidão igual a 1, o segundo pior recebe 2, e assim até que o melhor cromossomo tenha o valor de aptidão igual ao número que corresponde ao tamanho da população.
- Estado estacionário: diferente dos anteriores, esse método não é voltado particularmente à seleção de pais. Nesse caso, a ideia é garantir que uma nova população tenha uma grande parte de cromossomos que sobreviverão a próxima geração. A cada nova geração, alguns poucos bons cromossomos são selecionados para criação da descendência, enquanto cromossomos de baixa aptidão são removidos e novos são colocados em seus lugares. Todo o resto da população sobrevive para a próxima geração.
- Mutação: o operador de mutação altera aleatoriamente um ou mais genes de um cromossomo para aumentar a variabilidade da descendência criada pelo cruzamento. Isso permite ao GA encontrar soluções que contenham genes com valores não existentes na população inicial. O parâmetro que regula esse operador é chamado de *probabilidade de mutação*. Caso o operador de seleção reduza a diversidade na população, o operador de mutação eleva a diversidade novamente. Além disso, quanto maior a probabilidade de mutação, menor é a chance de convergência prematura. Entretanto, a alta probabilidade de mutação faz com que o GA torne-se um algoritmo de busca aleatória, segundo Dawid (2011). A Figura 3.2 ilustra um exemplo de mutação.

Nos algoritmos genéticos, a seleção dos indivíduos é dada por uma função denominada função de aptidão, responsável por atribuir um escore para cada cromossomo na população atual (Mitchell, 1995). A aptidão de um cromossomo depende de quão bem tal cromossomo resolve o problema. Em geral, os algoritmos genéticos são implementa-

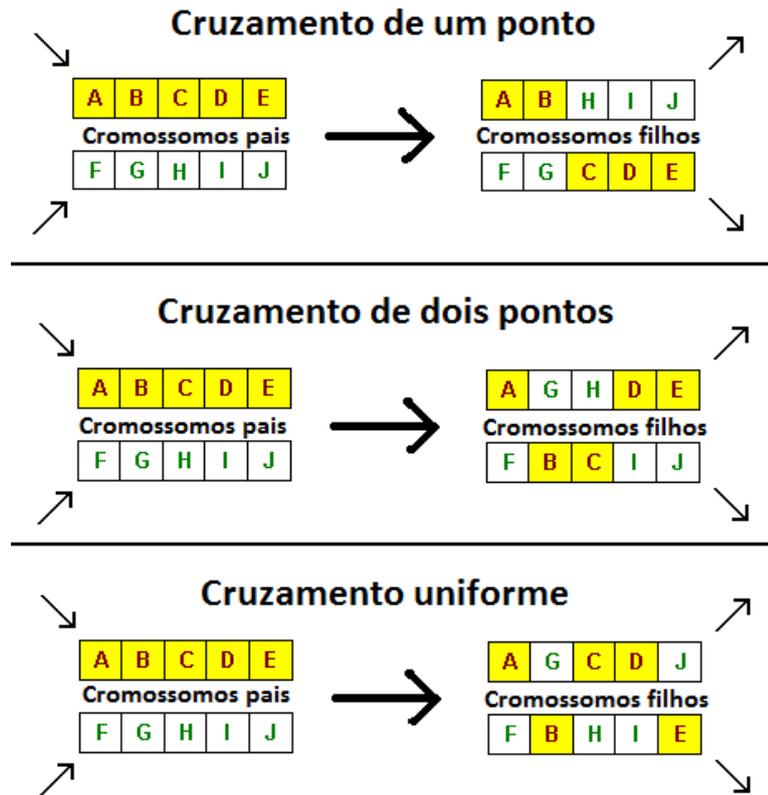


Figura 3.1: Exemplificação dos tipos de cruzamento implementados nos algoritmos genéticos. Fonte: (Silva, 2015)

dos visando à maximização dessa função. No entanto, quando o objetivo é minimizar tal valor, geralmente retorna-se o valor negativo que iria ser retornado pela função de aptidão. Por convenção, o melhor cromossomo de todas as gerações, isto é, que gera o melhor valor de aptidão, é chamado de cromossomo ótimo (Silva, 2015).

### 3.1.2 Algoritmo

Mitchell (1995) descreve os algoritmos genéticos de forma simples, como será detalhado a seguir, entretanto não é a única versão existente, apesar de ser a base para a maior parte das aplicações de GA, existem outras versões mais variações desse algoritmo.

1. Crie uma população aleatória P de N cromossomos (soluções candidatas para o

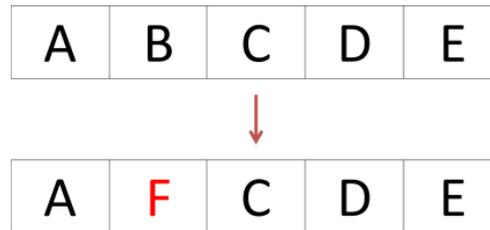


Figura 3.2: Exemplo da operação de mutação de cromossomos. Fonte: (Silva, 2015)

problema).

2. Avalie  $f(c)$  para cada cromossomo na população.
3. Gere uma nova população pela repetição dos seguintes passos até que a nova população atinja o tamanho  $N$ :
  - (a) Selecione dois cromossomos pais da população dando preferência aos de maior valor de aptidão. Automaticamente copie o melhor cromossomo para a próxima geração
  - (b) Dada a probabilidade de cruzamento, faça o cruzamento de dois cromossomos pais para gerar dois novos filhos. Se o cruzamento não foi executado, os cromossomos filhos são exatamente a copia dos pais.
  - (c) Copie o valor do cromossomo filho para a próxima geração.
4. Copie a nova geração criada sobre a população anterior.
5. Se a condição de término do *loop* é satisfeita, então pare e retorne a melhor solução na população atual. Caso contrário, vá para o passo (2).

### 3.1.3 Parâmetros

A utilização de algoritmos genéticos envolve a escolha de diversos fatores, entre eles, o número de cromossomos, o número de gerações, a probabilidade de mutação, a proba-

bilidade de cruzamento. A escolha desses fatores é crucial para que o algoritmo seja capaz de encontrar o resultado esperado.

A escolha do número de indivíduos é importante para a variabilidade da população. Quando maior o número de indivíduos, espera-se que maior seja a diferença entre eles, já que os  $N$  indivíduos iniciais, são criados de forma aleatória. No entanto, um elevado número de indivíduos pode piorar o desempenho desse algoritmo, visto que mais operações de cruzamento e mutações podem ser efetuados. Valores entre cem e mil têm sido usado frequentemente em trabalhos (Silva, 2015; Bäck e Schwefel, 1993; Behroozmand e Almasganj, 2007).

O número de gerações é crucial para garantir que o GA não convirja para algum mínimo local no espaço de resultados. Quando o número de gerações é consideravelmente baixo, a variabilidade da população pode não ser suficiente para que o algoritmo genético explore de forma mais satisfatória tal espaço. Por outro lado, quanto maior o número de gerações, maior o tempo de processamento necessário para a execução da tarefa. Embora não seja possível garantir que os algoritmos genéticos encontrem o melhor resultado (ótimo global) dentro do espaço de resultados, uma vez que ele nem sempre é totalmente conhecido, em geral, utiliza-se cem gerações por ser o suficiente para a convergência do resultado.

A probabilidade de cruzamento e mutação também são importantes na diversidade da população, visto que elas são responsáveis pela geração de novos indivíduos. Uma baixa probabilidade desses dois fatores pode impedir o algoritmo genético de explorar o espaço de resultados, convergindo prematuramente. Por outro lado, a alta probabilidade de mutação, como dito anteriormente, pode tornar o algoritmo genético um algoritmo de busca aleatória (Silva, 2015).

## 3.2 Otimização por Nuvens de Partículas

Nesta seção faz-se uma breve introdução à Inteligência de Enxames, seguida de uma visão geral da meta-heurística PSO, principal representante dos algoritmos baseados em Inteligência de Enxames.

### 3.2.1 PSO

Muitas espécies de animais se organizam em grupos nos quais geralmente há a presença de um líder que controla ou orienta o restante do grupo, como por exemplo o caso de um bando de leões ou matilha de lobos. Neste tipo de sociedade organizada, o comportamento dos indivíduos é fortemente ditado pela hierarquia social (Anderson e Franks, 2001). Mais curioso ainda é o comportamento das espécies que vivem em grupos em que nenhum líder pode ser identificado, por exemplo, um bando de aves ou um cardume de peixes. Nestes grupos sociais, os indivíduos não têm nenhum conhecimento do comportamento de todo o grupo e nem têm qualquer informação sobre o meio ambiente. Apesar disso, eles têm a capacidade de juntar-se e se moverem juntos, com base em interações locais entre indivíduos, ou seja, um comportamento social complexo surge a partir de interações simples entre indivíduos de um mesmo grupo (Coelho Filho et al., 2013).

Com base no comportamento descrito acima, conhecido como inteligência de enxames, Eberhart e Kennedy (1995) desenvolveram o algoritmo conhecido como otimização por nuvens de partículas. Em analogia aos paradigmas de computação evolucionária, um enxame é semelhante a uma população, enquanto que uma partícula é semelhante a um indivíduo.

Simplificando, as partículas percorrem um espaço de busca multidimensional, em que a posição de cada partícula é ajustada a cada iteração do algoritmo de acordo com a sua própria experiência e a de seus vizinhos. Cada partícula é uma solução viável

para o problema e tem sua qualidade ou valor dado por uma função de aptidão, tal como os algoritmos genéticos. A Figura 3.3 ilustra o espaço de busca percorrido pelas partículas.

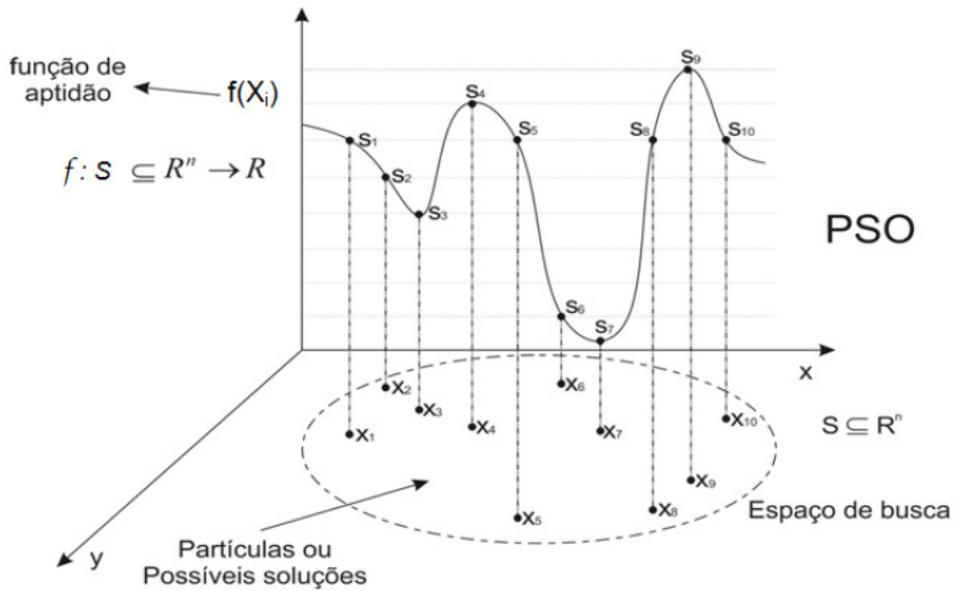


Figura 3.3: Partículas em um espaço multidimensional. Fonte: (Coelho Filho et al., 2013)

Para um enxame de  $N$  partículas, a velocidade da  $i$ -ésima partícula na posição  $X_i(t)$  é calculada pela seguinte equação:

$$V_i(t+1) = \underbrace{w_i(t) \cdot V_i(t)}_{\text{Velocidade prévia}} + \underbrace{c_1 \cdot r_1 \cdot (Pbest_i(t) - X_i(t))}_{\text{Componente cognitivo}} + \underbrace{c_2 \cdot r_2 \cdot (Gbest - X_i(t))}_{\text{Componente social}} \quad (3.1)$$

Onde  $c_1$  e  $c_2$  são constantes de aceleração,  $r_1$  e  $r_2$  são variáveis aleatórias e  $pbest_i(t)$  e  $gbest$  são respectivamente a melhor posição que a partícula  $i$  já ocupou e a melhor posição já ocupada dentre todas as partículas até o momento atual. A variável  $w$ , é chamada de peso inercial e serve como uma espécie de freio que diminui gradativamente a velocidade da partícula a cada iteração. A nova posição da partícula é dada por:

$$\underbrace{X_i(t+1)}_{\text{Posição atual}} = \underbrace{X_i(t)}_{\text{Posição anterior}} + \underbrace{V_i(t+1)}_{\text{Velocidade atual}} \quad (3.2)$$

A Figura 3.4 segue a ilustração sobre a movimentação de uma partícula em um espaço bidimensional.

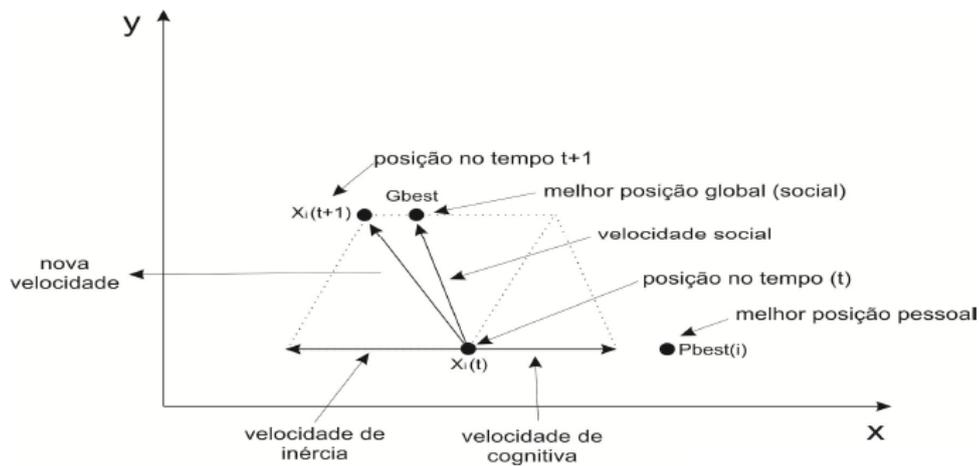


Figura 3.4: Movimentação de uma partícula. Fonte: (Coelho Filho et al., 2013)

O vetor velocidade dirige todo o processo de otimização e reflete ambos os conhecimentos da partícula: o conhecimento advindo da própria experiência e o adquirido com a troca de informações com sua vizinhança. O conhecimento individual adquirido da sua própria experiência é referenciado como o componente cognitivo da partícula proporciona a informação sobre a menor distância da solução ideal que aquela partícula

já esteve. Esta posição é denominada como *melhor posição pessoal* (*pbest*) da partícula e influencia somente no seu próprio movimento. A informação obtida da troca de informações com a vizinhança é referenciada como o componente social na equação de velocidade. O componente social é também conhecido como *melhor posição global* (*gbest*) e influencia no movimento de todas as partículas do enxame.

O PSO realiza um processo de otimização iterativo. A cada iteração a velocidade e posição de cada partícula em cada dimensão é atualizada e depois disso é calculado o valor da aptidão de cada partícula, de acordo com a função de avaliação ou função objetivo da otimização. Com base na aptidão da partícula são ajustadas as posições das partículas *Gbest* e *Pbest*. Portanto a cada iteração um total de  $n$  funções de avaliação é calculado onde  $n$  é o número de partículas do enxame. A função de aptidão é dependente do problema de otimização, que pode ser para maximização ou minimização do valor da função objetivo (Coelho Filho et al., 2013; Queiroga, 2016).

O PSO é descrito em alto nível pelo pseudocódigo abaixo e pelo diagrama de fluxo na Figura 3.5. Na versão pura, o PSO inicia com vetores de velocidade nulos. Por isso, nesta iteração, o *gbest* possui maior poder de influência do que nas iterações seguintes.

1. Inicializa a nuvem de partículas com posições aleatórias, e a função objetivo de cada partícula (*pbest inicial*) e encontra o *gbest* inicial;
2. Calcula-se os vetores de velocidade e posições através das equações 3.1 e 3.2.
3. O *pbest* é atualizado para cada partícula individualmente se houver melhora em sua busca pela solução de objetivo. E o *gbest* é atualizado, caso alguma partícula encontre uma solução melhor do que o melhor valor já encontrado.
4. Se os critérios de término são satisfeitos, retorne o *gbest*. Caso contrário, vá para o passo 2.

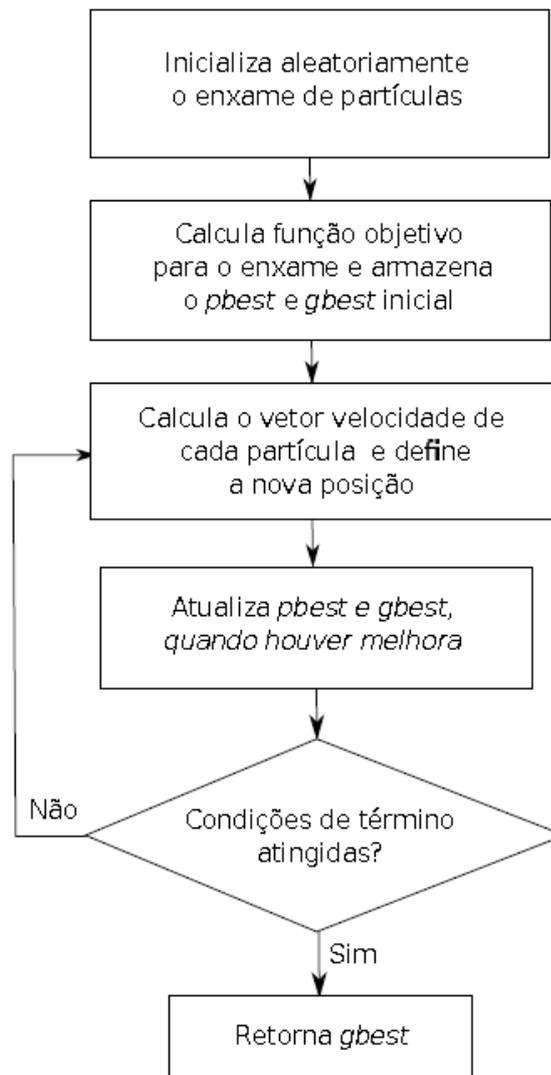


Figura 3.5: Diagrama de Fluxo do PSO básico. Fonte: (Queiroga, 2016)

## Capítulo 4

# Trabalhos Relacionados

Esse capítulo revisa trabalhos consolidados na literatura na área de algoritmos de casamento de impressões digitais, além de trabalhos que usam algoritmos evolucionários para otimização de parâmetros e união de algoritmos diferentes.

O VeriFinger é um kit de desenvolvimento (do inglês: *Standard Development Kit*, SDK) para impressão digital que é usado em mais de 100 países, e tem sido comercializado há mais de 16 anos. Este SDK foi premiado no *Fingerprint Verification Competition* (FVC) - competição que testa a velocidade e confiabilidade de SDK's - na área de impressão digital, e não foi superado desde o ano 2000. O VeriFinger foi desenvolvido pela Neurotechnology, que é uma empresa lituana que atua em pesquisa na área de biometria há mais de 25 anos, essa empresa desenvolve pesquisas em impressões digitais, reconhecimento facial, reconhecimento por voz, e por íris. A Neurotechnology fornece seus serviços para fins comerciais e é bastante conceituada no âmbito internacional (Neurotech, 2014). Apesar de ser considerado um algoritmo de fins comerciais e não haver publicações a respeito do processo de desenvolvimento desse algoritmo. O Verifinger conta com uma taxa de acerto de 99,8% no projeto FVC On Going (BioLab, 2006).

O Minutia Cylinder-Code (MCC), de acordo com vários estudos (Cappelli et al.,

2010b), é um dos mais precisos algoritmos de casamento local baseados em minúcias. O MCC utiliza raio fixo para busca local de minúcias, isso o torna mais resistente à ausência de informações e a ruídos. Também é tolerante à distorções locais e erros na extração de informações e regiões ruidosas, onde se encontram informações espúrias. O algoritmo possui implementação simples, rápida e eficaz. Esse algoritmo foi desenvolvido por pesquisadores da Universidade de Bolonha, Itália, e em sua versão 1.4 conta com uma taxa de acerto de aproximadamente 99.6% (BioLab, 2006).

Medina-Pérez et al. (2016) propõem um método de identificação de *clusters* de minúcias após uma etapa de casamento local. O algoritmo sugere que um conjunto de *clusters* de minúcias dissociados e com alguma interseção têm maior capacidade de identificar regiões de interesse do que cada um isoladamente. Esse algoritmo pode, portanto, ser usado como auxiliar a algoritmos já existentes.

O M3gl (Medina-Pérez et al., 2011) é um algoritmo de casamento de impressões digitais que segue o padrão ISO/IEC 19794-2:2005 (Committees, 2005) e se baseia em conjuntos triangulares de minúcias, chamados *triplets*, como estrutura básica. A partir de dois conjuntos de *triplets*, um em cada impressão digital, inicia-se a busca por estruturas correspondentes. Essa correspondência é dada pela comparação entre as distâncias euclidianas e a relação entre o ângulo entre cada minúcia dos triângulos. Após a etapa local de casamento entre as minúcias, é realizada uma sobreposição do conjunto completo de minúcias, onde todo o conjunto é rotacionado baseado no melhor casamento da etapa anterior. Este algoritmo, é considerado bastante rápido, simples e eficaz, apresentando taxa de acerto superior a 99% no FVC (BioLab, 2006).

O SourceAFIS é um sistema de identificação por impressão digital (AFIS) de código aberto. Ele essencialmente compara duas impressões digitais e decide quando elas pertencem a mesma pessoa. O seu algoritmo é baseado em expansão de grafos baseado em vizinhos mais próximos. Além disso a API está disponível em .NET e parcialmente em Java. O SDK acompanha um *software* de visualização e análise das impressões digitais

que permite a exibição das etapas de extração de características das imagens, além de apontar as informações que estão presentes em ambas as imagens que representa o grau de similaridade indicada para a aceitação da comparação. O SourceAFIS é um projeto pessoal desenvolvido por Robert Važan que disponibilizou o código e posteriormente conseguiu contribuições de outros interessados no projeto (Važan, 2009). O projeto inicial obteve uma taxa de acerto de 98,8% segundo o FVC (BioLab, 2006).

O trabalho de Silva (2015) realiza a otimização de parâmetros de traços multibiométricos de Geometria Palmar e Impressão Palmar utilizando Algoritmos Genéticos. No trabalho utilizou-se a base multibiométrica de imagens BioPass-UFPB. Ao todo utilizaram-se 99 atributos sendo 85 geométricos e 14 de textura, de cada imagem. Os resultados mostraram que, nas duas abordagens de verificação empregadas, os algoritmos genéticos conseguiram melhorias superiores a 30% e 90% da taxa EER em relação ao caso em que o GA não foi aplicado. Na classificação, o uso de algoritmos genéticos conseguiu reduzir na média o número de templates a serem recuperados pelo sistema para garantir que ao menos um desses seja da mesma classe da amostra de referência.

Dalila et al. (2015) propõem o uso de uma abordagem híbrida utilizando os Algoritmos Genéticos e a Otimização por Nuvens de Partículas para otimização de atributos de pontuação de abordagens multibiométricas. O foco daquele trabalho é encontrar os pesos ótimos para a fusão de todas as modalidades biométricas. Dalila et al. chegam a conclusão de que não há muita diferença ao fim da otimização entre GA, PSO e GA-PSO (técnica proposta) para os resultados visto que todos encontraram as melhores combinações de parâmetros. Porém, é possível observar diferença no tempo de processamento, que é onde a técnica proposta de um algoritmo de otimização híbrido ganha, sendo adicionado o benefício de busca rápida que é propriedade do PSO e ainda assim cobrindo um grande espaço de busca de forma eficiente. O método conseguiu alcançar o ótimo global em apenas duas iterações enquanto o PSO alcançou em 9 e o GA em apenas 38 iterações.

Brahnam et al. (2013) apresentam um sistema de casamento combinado para verificação de impressões digitais utilizando descritores baseados em minúcias e em correlação, além de textura da área próxima à minúcia a fim de extrair uma grande quantidade de características que a descrevam. O algoritmo foi testado nos quatro bancos de imagens do FVC2004 e a avaliação foi realizada através da análise do valor EER. Os experimentos realizados mostraram que a abordagem tem bom desempenho em todos os bancos testados mesmo sem treinamento para os diferentes bancos de dados testados e obteve taxas de acerto semelhantes a sistemas comerciais que competiram no FVC 2004.

# Capítulo 5

## Materiais e Métodos

Este Capítulo descreve os recursos utilizados e os métodos desenvolvidos. Inicialmente, apresenta-se o ambiente desenvolvido, em seguida o método proposto é apresentado juntamente com a metodologia de treinamento e validação.

### 5.1 Ambiente de Desenvolvimento

O algoritmo de casamento de impressões digitais foi desenvolvido na linguagem de programação C++ versão 11.0, no ambiente de desenvolvimento *Microsoft® Visual Studio Enterprise 2013*. Para realizar testes e validação do algoritmo foi necessário um ambiente de apoio para visualização de desempenho e gráfica do algoritmo, para isso foram desenvolvidos aplicativos utilizando a linguagem C# no mesmo ambiente de desenvolvimento com o objetivo de facilitar esse processo. Por fim, os treinamentos utilizando algoritmos bioinspirados foram realizados utilizando a linguagem de programação Java em sua versão 1.8 64 *bits* no ambiente de desenvolvimento *NetBeans®* em sua versão 8.0.2. A utilização dos Algoritmos Genéticos neste último ambiente de programação foi realizada pela ferramenta JGAP em sua versão 3.6.3 (Meffert et al., 2012). O algoritmo de Otimização por Nuvens de Partículas de Partículas foi implementado em sua versão

mais simples, seguindo o algoritmo descrito no Capítulo 2.

Para realizar os testes e treinamento do algoritmo proposto, foi utilizado um computador modelo *Desktop* com CPU intel® Core i7-4770 3,4 GHz, 24 GB de memória RAM 1600 MHz, com disco de armazenamento modelo SSD. Sistema Operacional *Microsoft® Windows 10 Pro 64 bits*.

## 5.2 Método Proposto

Este trabalho propõe uma nova abordagem de casamento de impressões digitais. O método proposto foi criado a partir de vários algoritmos disponíveis na literatura com a finalidade de juntar suas características fortes e tornar o algoritmo proposto mais robusto.

O desenvolvimento deste método se divide em três etapas: Casamento em escopo local; Casamento em escopo global e treinamento do algoritmo utilizando algoritmos bioinspirados. Conforme ilustrado na Figura 5.1

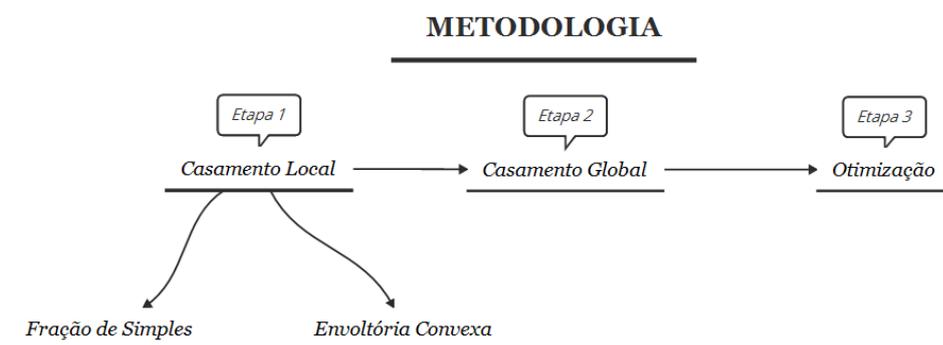


Figura 5.1: *Workflow* do desenvolvimento do método proposto.

### 5.2.1 Casamento Local

O casamento local utiliza informações das minúcias em um contexto isolado. Nessa metodologia as minúcias são avaliadas como estruturas isoladas onde não se tem conhecimento das demais estruturas existentes no dedo. Dessa forma, a partir de um escopo local, o casamento de impressões digitais avalia o relacionamento entre as minúcias e suas vizinhas, utilizando alguns critérios de similaridade. Assim, se o relacionamento de uma minúcia com sua vizinha for semelhante a outro entre duas minúcias presentes na amostra de comparação, entende-se que as estruturas podem ser equivalentes.

Em outras palavras, considerando que uma minúcia pode ser descrita como uma tupla  $M_i = (x_i, y_i, \theta)$ , onde  $(x_i, y_i)$  representam a posição da minúcia e  $\theta$  representa a direção da minúcia, em um intervalo  $[0, 2\pi]$ , o casamento local, que é baseado em expansão de grafos, parte da seguinte ideia: dada duas minúcias  $M_i$  e  $M_j$ , a minúcia  $M_i$  na impressão digital  $A$  corresponde à minúcia  $M_j$  na impressão digital  $B$  se o relacionamento com o conjunto de vizinhas de  $M_i$  for semelhante ao relacionamento entre a minúcia  $M_j$  e suas vizinhas. Alguns algoritmos, no entanto, tratam a vizinhança como uma região de raio  $r$  ao redor das minúcias.

Tanto a abordagem baseada em vizinhas mais próximas quanto a abordagem baseada em busca de vizinhos pela área ao redor das minúcias possuem desvantagens. A abordagem baseada em vizinhos mais próximos pode ser limitada se a vizinhança estiver em uma região com muito ruído, onde os vizinhos encontrados são falsas minúcias, ou ainda quando a minúcia não for encontrada, como pode ser visto na Figura 5.2. Por outro lado a região limitada por raio pode encontrar poucos vizinhos, ou os vizinhos relevantes podem estar próximo à fronteira do raio delimitado, mas fora dela, sendo ignorados como pode ser visto na Figura 5.3.

O método proposto utiliza o método baseado em expansão de grafos avaliando os vizinhos mais próximo das minúcias. E como afirmam Cappelli et al. (2012), os atributos das minúcias não podem ser comparados diretamente, pois a posição da minúcia em uma

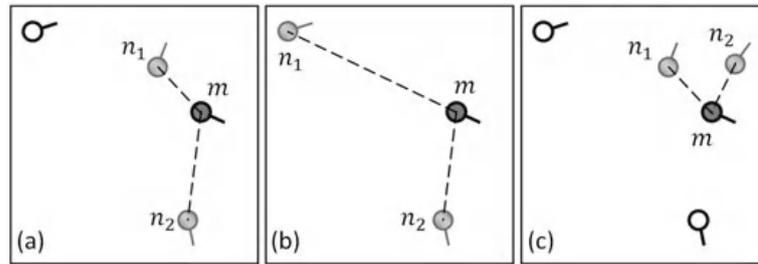


Figura 5.2: Um exemplo de busca local baseada em vizinhos mais próximos para  $K = 2$ . Erros ocorrem devido à minúcias ausentes(b) e informações espúrias(c)

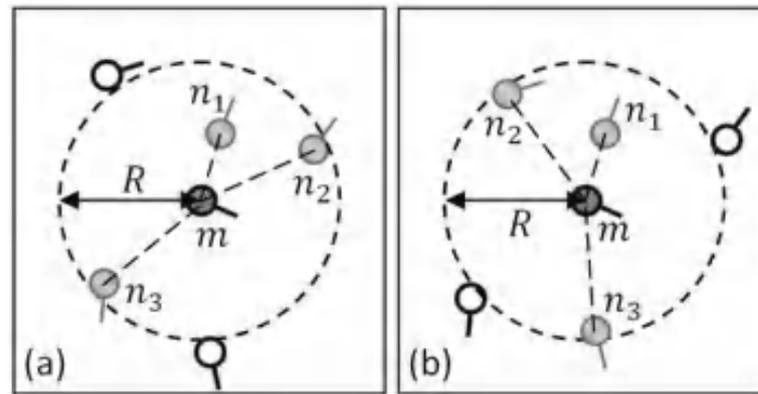


Figura 5.3: Um exemplo de raio fixo para busca de minúcias na vizinhança. (a) e (b) demonstram o problema decorrente da área de fronteira

captura de digital dificilmente será a mesma posição em uma outra captura da mesma digital, as relações entre as minúcias, entretanto, permanecem as mesmas, toleradas algumas pequenas variações devido a variações inerentes à captura. O método proposto utiliza estruturas de relacionamento entre minúcias de uma mesma vizinhança, de modo que esse relacionamento possa ser classificado como único em uma busca na impressão digital.

Para iniciar a comparação entre as amostras de impressão digital, um conjunto de segmentos que ligam as minúcias às suas vizinhas é criado. A avaliação será baseada nas características presentes nesses relacionamentos, pois, por mais que uma minúcia possa ter sua posição e ângulo de direção variados nas diferentes capturas, o relacionamento entre duas minúcias permanecerá inalterado.

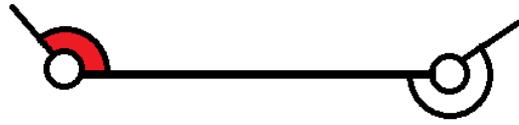


Figura 5.4: Informações do Segmento: *Reference Angle*



Figura 5.5: Informações do Segmento: *Neighbor Angle*

Assim, dado um conjunto de minúcias  $M_i$  do *template* A, para cada minúcia, o algoritmo buscará os  $k$  vizinhos mais próximos e criará segmentos entre  $M_i$  e estes. Os segmentos geralmente representam uma informação única de relacionamento entre duas minúcias e é mais confiável do que as informações delas isoladamente. Partindo do tamanho e das informações de angulação do segmento pode-se extrair as seguintes propriedades:

- *Reference Angle*: O valor do ângulo formado entre a minúcia de referência, de onde o segmento inicia, e o ângulo do segmento, no sentido horário (Figura 5.4).
- *Neighbor Angle*: O valor do ângulo formado entre a minúcia vizinha, onde o segmento chega, e o ângulo do segmento no sentido anti-horário (Figura 5.5).
- Tamanho do Segmento: Distância euclidiana entre as duas minúcias (Figura 5.6).

As relações entre a direção das minúcias e o ângulo do segmento são importantes para aumentar a confiabilidade no casamento, ele as utiliza com o objetivo de ignorar



Figura 5.6: Informações do Segmento: Tamanho do segmento

alterações globais de ângulo decorrente de diferentes capturas, visto que o ângulo de cada minúcia pode ser alterado em conjunto, mas as relações entre elas obedecem um padrão. Esse algoritmo não armazena nenhuma informação sobre a posição da minúcia. Isso acontece porque ele as ignora no padrão encontrado. Após a construção de todos os segmentos, toda a informação é armazenada em uma matriz  $M[n,k]$ , onde  $n$  é a quantidade de minúcias presentes no *template*. Em um segundo passo do algoritmo, cada segmento gerado no *template* B procurará um segmento correspondente na matriz do *template* A. A comparação entre dois segmentos dá-se pela subtração dos valores do tamanho dos segmentos, *Reference Angles* e *Neighbors Angles*. Os segmentos são ditos casados se esses valores forem menores do que os limiares de ângulo e tamanho pré-definidos.

Quando um vetor do *template* A corresponde a um encontrado no *template* B, o algoritmo entende que as minúcias de referência são correspondentes (ver Figura 5.7). Baseado nessa premissa, é possível que as minúcias na vizinhança também possam encontrar suas correspondentes pois essa região, agora, é candidata para ser uma região de interseção - A Figura 5.8 mostra um exemplo de região de interseção. Então a minúcia pareada no *template* B procura entre seus segmentos vizinhos se algum deles corresponde aos segmentos vizinhos da minúcia pareada no *template* A.

Quando os vizinhos casam, o ciclo de busca por outros vizinhos reinicia. Cada novo par de minúcias casadas procura entre seus segmentos vizinhos por segmentos correspondentes. Esse ciclo recursivo continua até que o grafo atinja as bordas da região de interseção, onde não haverá mais minúcias correspondentes nos dois *templates*. O grafo tende a ficar significativamente menor em pares de impressões digitais de indivíduos diferentes do que em pares do mesmo indivíduo.

Em uma etapa final, a lista de minúcias pareadas entre o *template* A e B é formada. A Figura 5.9 mostra o melhor grafo gerado a partir do casamento de dois *templates*. Um grande número de informação pode ser inferida dessa comparação, tal como quantidade



Figura 5.7: Exemplo de minúcias casadas



Figura 5.8: Exemplo de região de interseção



Figura 5.9: Exemplo de um grafo completo

de minúcias casadas do total de minúcias de cada *template*; quantidade de minúcias de mesmo tipo que foram pareadas; quantidade de segmentos presentes no grafo, entre outras. Essas informações podem ser usadas para estabelecer uma pontuação de similaridade entre as amostras.

### 5.2.2 Aplicação do Método Envoltória Convexa

Uma medida importante utilizada no momento de comparação de segmentos é a *fração de casamento*. Essa medida indica a quantidade de minúcias que casaram em relação ao total de minúcias presente nos templates. Pode ser expressa pela Equação 5.1.

$$MatchFraction = \frac{MatchedMinutiae}{\frac{MinutiaA + MinutiaB}{2}} \quad (5.1)$$

Onde *MatchFraction* indica o cálculo realizado para a fração de casamento; *MinutiaA* e *MinutiaB* indicam a quantidade de minúcias presentes nos *templates* A e B

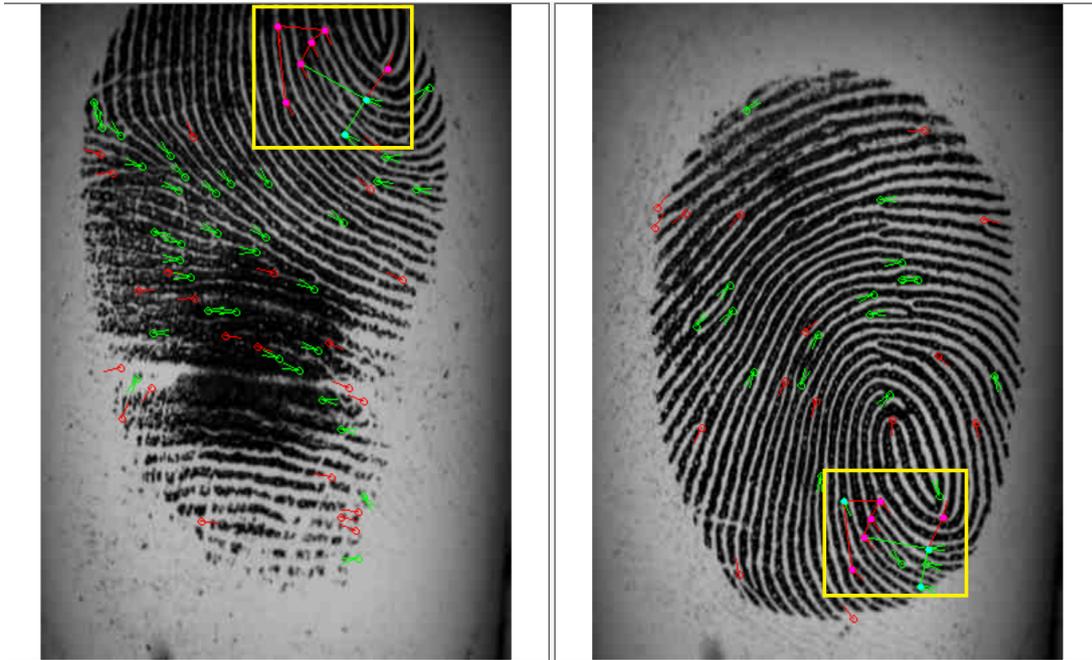


Figura 5.10: Possível região de intersecção entre duas imagens.

respectivamente, e *MatchedMinutiae* representa a quantidade de minúcias pareadas no fim da comparação.

A *fração de casamento* é usada por alguns autores (Maltoni et al., 2009) como medida única para definição de similaridade entre duas amostras (Equação 5.1).

Neste trabalho, diferentemente dos métodos de pontuação sugeridos por Maltoni et al. (2009) e Medina-Pérez et al. (2011), percebeu-se que esta medida, com a finalidade de indicar similaridade entre duas amostras, pode falhar em casos em que há a captura parcial das duas amostras. Quando duas amostras têm, apenas, parte distintas da impressão digital capturadas, a área de intersecção entre as duas digitais diminui, como pode ser vista na Imagem 5.10. Dessa forma a quantidade de minúcias na região de intersecção é muito menor do que a área total da impressão digital.

O função envoltória convexa foi adotada neste caso para que a área avaliada se restrinja à região de intersecção das duas amostras. Dessa vez, a *fração de casamento* conterà apenas as minúcias presentes nessa área de intersecção. Ou seja, será a fração

entre as minúcias casadas e as minúcias que possivelmente casariam por estarem na área abrangida pelas duas amostras. Conforme segue:

$$MatchFraction = \frac{MatchedMinutiae}{\frac{MinutiaA' + MinutiaB'}{2}} \quad (5.2)$$

Onde, dessa vez,  $MinutiaA'$  e  $MinutiaB'$  são a quantidade de minúcias de A e de B, respectivamente, que estão em uma possível região de interseção entre as duas amostras.

No caso da Figura 5.10, onde a fórmula 5.1 resultaria na divisão entre as 9 pares de casamento entre os templates e o a média de 53 minúcias nos dois templates; aplicando a fórmula 5.2 divide-se as 9 minúcias pareadas por um total de 11 minúcias que estão na possível área de interseção. Essa diferença entre as fórmulas pode determinar se as amostras serão aceitas ou falsamente rejeitadas visto que são amostras que pertencem ao mesmo dedo (Figura 5.11).

Neste método, informações extraídas a partir do grafo formado após a execução do algoritmo são utilizadas para a formação de uma pontuação que indicará o grau de similaridade entre as amostras, entre essas informações a fração de casamento é utilizada.

### 5.2.3 Casamento Global

A abordagem de casamento global de impressões digitais é uma metodologia diferente do casamento local e tem vantagens e falhas também distintas. O objetivo do uso desse algoritmo é justamente por essa variabilidade de acertos e falhas. Com o uso de mais de um método de casamento espera-se obter acertos onde a abordagem de casamento local falha, de forma que possam se complementar.

O casamento global funciona de maneira mais simples. O objetivo aqui é alinhar as duas digitais de forma que encontre-se a região de intersecção a partir da sobreposição das amostras.

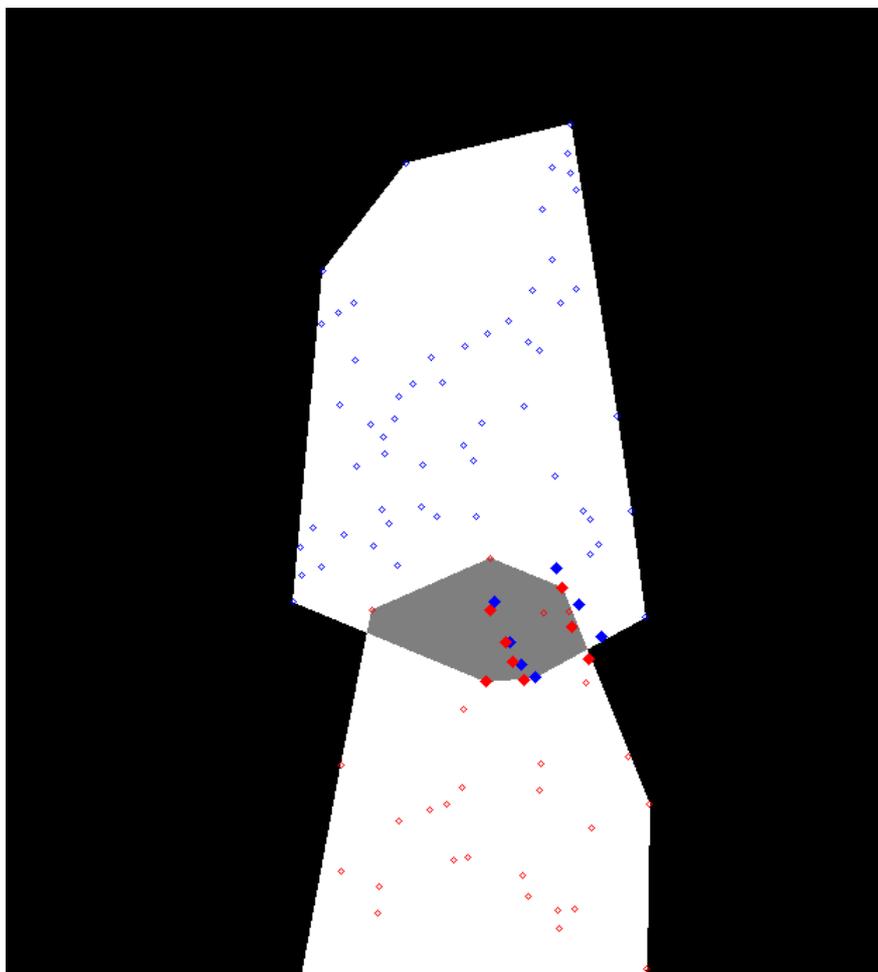


Figura 5.11: Imagens sobrepostas indicando possível região de interseção. As circunferências azuis indicam minúcias não casadas que pertencem à amostra A, as circunferências vermelhas representam amostras não casadas pertencentes à amostra B e os círculos azuis e vermelhos representam os pares casados das amostras A e B respectivamente.

A vantagem dessa abordagem é que as impressões digitais são analisadas como um todo e esse método é menos afetado por imagens de baixa qualidade ou regiões ruidosas. A sua desvantagem, entretanto, é que este método não é robusto às deformações não lineares da pele que são apresentadas no momento da captura das imagens. Observa-se, por sua vez, que o casamento local tem a dificuldade de ultrapassar fronteiras de ruídos que podem ser vencidas por esse método e é capaz de lidar com as distorções não lineares apresentadas pela pele.

Essa abordagem utiliza informações do casamento local para o alinhamento entre as amostras. Como a abordagem local constrói um grafo de similaridade correspondente nas duas imagens, é possível analisar as minúcias que foram pareadas de forma a estabelecer um possível alinhamento entre as amostras e tentar identificar outras minúcias que não foram encontradas pelo casamento local.

Assim o casamento global utiliza as minúcias do grafo gerado pela abordagem local da seguinte forma:

1. Alinhamento: Nesta etapa busca-se o alinhamento das amostras a partir de informações obtidas no casamento local. Como já existem informações a respeito das minúcias correspondentes da etapa anterior, identifica-se a possível diferença de direção entre as duas impressões digitais e é realizado um processo de rotação e translação, onde todas as minúcias de uma das digitais são rotacionadas e transladadas, tendo como o ponto central a minúcia selecionada para alinhamento. O processo de alinhamento é realizado com a finalidade de colocar as minúcias dos *templates* no mesmo plano para que elas possam ser avaliadas. Esse processo é realizado até que todas as minúcias previamente pareada no casamento anterior sejam avaliadas a fim de se encontrar o melhor alinhamento entre as duas digitais.
2. Comparação: Após cada tentativa de alinhamento, as minúcias das amostras serão avaliadas a fim de encontrar uma correspondente na amostra que está sendo

comparada.

A avaliação acontece calculando a proximidade entre as minúcias. Assim as minúcias do template A são analisadas e verifica-se se há alguma minúcia do template B próxima a ela. Assim se uma minúcia  $M_b$  pertencente ao *template* B estiver a uma distância euclideana menor que um dado  $r$  e a direção da minúcia avaliada também for menor que um valor de aceitação, pode-se considerar que as minúcias são correspondentes.

O processo de comparação é semelhante ao apresentado por Maltoni et al. (2009) na Figura 5.12. Se existirem mais de uma minúcias dentro da região de tolerância, a de menor distância será priorizada.

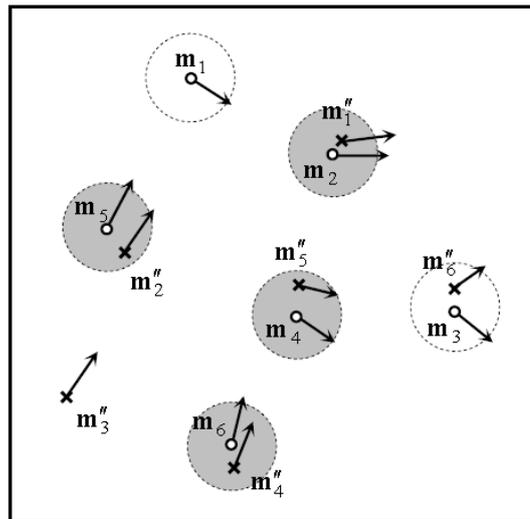


Figura 5.12: Avaliação de minúcias no casamento Global.

#### 5.2.4 Treinamento

O treinamento consistiu em usar algoritmos bioinspirados para ajustar os parâmetros de pontuação das variáveis geradas a partir do método de comparação local. O objetivo é encontrar o conjunto de pesos otimizados que ajusta o algoritmo para que se obtenha a

menor taxa de erro nos bancos de dados usados para treinamento, bem como identificar o melhor método de treinamento. Assim, com o auxílio dos algoritmos bioinspirados, pesos serão atribuídos às variáveis encontradas pelo casamento local e a função objetivo será a diminuição da taxa de erro nos bancos de dados treinados. No total foram otimizados 11 parâmetros na primeira etapa de treinamento, estes faziam parte do cálculo de pontuação do casamento local; e na segunda parte do treinamento foram otimizados três variáveis, a soma ponderada entre os valores da pontuação do casamento local, global, e fração do casamento das minúcias pareadas no segundo casamento; totalizando 14 variáveis treinadas.

A função objetivo deste método recuperou todas as instâncias de casamento de impressões digitais que são realizadas nos bancos de dados atribuindo-lhes o conjunto de parâmetros que eram escolhidos a cada execução destes métodos. Dessa forma, cada instância de execução dos algoritmos de otimização implicou em um recálculo em todas as instâncias de casamentos que eram executada nos bancos de treinamento. Assim, a melhor função de aptidão encontrada foi aquela cujo conjunto de parâmetros encontrado pelos algoritmos de otimização geravam menor EER no banco de treinamento.

Após o treinamento do casamento local, como o casamento global não necessitava de treinamento de parâmetros foi realizado outro treinamento para atribuir pesos ao casamento local e global para assim criar uma pontuação para o casamento entre as impressões digitais. Dessa forma os algoritmos bioinspirados foram executados duas vezes, uma em cada etapa, a primeira para definir valores utilizados no casamento local, e a segunda para atribuir pesos ao agrupamento de algoritmos.

Utilizaram-se os algoritmos de otimização por nuvens de partículas e algoritmos genéticos. Para ambos o número de partículas, ou de cromossomos nos algoritmos genéticos, que indicam o tamanho da população, foi 100. E a quantidade de gerações nos algoritmos genéticos e de épocas no método de otimização por nuvens de partículas foi 20. No método de otimização por nuvens de partículas, como não contou com apoio

de nenhuma ferramenta externa, como aconteceu com o GA, pôde-se ter um maior controle e foi ajustado para que, caso a melhor resposta encontrada persistisse por 10 gerações, o algoritmo encerrar.

Nos algoritmos genéticos foi utilizado o método de *cruzamento em um ponto*, com probabilidade de 35%. A probabilidade usada para a taxa de mutação foi de 12% e o método de seleção utilizado foi o estacionário. O conjunto de parâmetros utilizados são o de configuração padrão da ferramenta JGaP (Meffert et al., 2012).

Já o método de otimização por nuvens de partículas, teve o mesmo conjunto de atributos utilizados no trabalho de Duraisamy e Kayalvizhi (2010). O *Peso Inercial* máximo utilizado foi de 0.4 variando até 0.1 no fim da execução do algoritmo. O *componente cognitivo local* teve seu peso igual a 2, do mesmo modo que o *componente cognitivo global*. A *probabilidade de morte* da partícula utilizada foi de 10%.

### 5.2.5 Bancos de Dados Utilizados

O treinamento foi realizado utilizando validação cruzada (método *leave-one-out*) com seis bancos de dados de versões diferentes do projeto FVC (BioLab, 2006). Foram escolhidos bancos de dados de impressões digitais coletadas a partir de sensores óticos com resolução de 500dpi ou muito próximo a essa. São eles: Banco DB3 do FVC 2000, que contém 800 imagens, sendo 80 dedos e 10 amostras por dedo; Banco DB1 do banco 2002, que contém 800 imagens, sendo 80 dedos e 10 amostras por dedo; Bancos DB1 e DB2 sendo 800 imagens, sendo 80 dedos com 10 amostras por dedo, cada; o Banco DB2 do FVC 2006, com 1680 imagens, sendo 140 dedos com 12 amostras cada e o Banco Multiobiométrico BioPass-UFPB, que contém um total de 720 imagens, sendo 90 dedos com 8 amostras por dedo (Silva, 2015). No total foram realizados seis treinamentos na validação cruzada.

O banco de dados utilizado do projeto FVC 2006, originalmente possui imagens com resolução de 569dpi, entretanto as imagens foram redimensionadas para que a

resolução chegasse a 500dpi e se igualassem às demais imagens utilizadas nos testes deste trabalho.

Como a etapa de casamento não trabalha diretamente sobre as imagens digitais e sim atua sobre os *templates*, as características das imagens foram extraídas utilizando o mesmo algoritmo de extração, a fim de manter a igualdade entre as comparações.

### 5.2.6 Validação do Método Proposto

Para validar o método proposto, é necessário dividir o conjunto de testes em duas etapas. A primeira valida a otimização realizada pelos algoritmos bioinspirados. Aqui se verifica o desempenho no algoritmo utilizando pesos unitários e o compara com seu desempenho utilizando pesos otimizados para suas variáveis.

Na segunda etapa o algoritmo, já otimizado, é comparado com os outros algoritmos disponíveis no mercado, o algoritmo considerado estado da arte, o MCC, que tem sua *SDK* disponibilizada para fins de pesquisa científica (Cappelli et al., 2010a, 2011; Ferrara et al., 2012) e outros dois algoritmos de casamento de impressões digitais de código aberto, o M3GL e o SourceAFIS, que também participaram do projeto FVC OnGoing. Além disso, são utilizadas informações presentes nas versões anteriores do projeto BioLab (2006) para verificar o desempenho que o algoritmo alcançaria nas versões anteriores daquele projeto. Esta etapa é realizada com o objetivo de verificar se o algoritmo proposto é capaz de competir com os algoritmos no estado da arte e os algoritmos comerciais.

São realizados testes *um-para-um* utilizando o protocolo do FVC para verificação da taxa de erro igual (EER).

## Capítulo 6

# Resultados e Discussão

Neste capítulo serão validados os métodos propostos. Inicialmente, o algoritmo será executado com parâmetros em peso unitário. Em seguida, os parâmetros serão otimizados utilizando os algoritmos bioinspirados. Analisa-se-á em seguida o desempenho dos algoritmos de otimização. Por fim o método proposto será comparado com resultados de outros projetos com seus desempenhos no projeto FVC e com o algoritmo considerado o estado da arte na área de casamento de impressões digitais, o *Minutia Cylinder-Code*.

Para fins de concisão, as bases de dados utilizadas neste estudo e já descritas no capítulo que trata os materiais usados serão referidas nas tabelas da seguinte forma: base de dados DB2 do projeto FVC 2006, como DB06; a base de dados DB1 do projeto FVC 2004 como DB04-1 e base DB2 da mesma versão do projeto como DB04-2; da mesma forma para as respectivas bases de dados para as versões 2002 e 2000 deste projeto. A base do projeto Biopass-UFPB será referida pelo nome BioPass.

## 6.1 Desempenho do método proposto utilizando pesos unitários

A Tabela 6.1 descreve a taxa de erro do algoritmo proposto utilizando pesos unitários em seus atributos que compõem a pontuação de similaridade.

Tabela 6.1: Casamento local utilizando pesos unitários nas variáveis que compõem a pontuação de similaridade.

	DB06	BioPass	DB04-2	DB04-1	DB02	DB00	Média
EER	1,05%	3,17%	5,92%	6,60%	4,84%	3,71%	4,22%

A otimização mostra-se necessária porque podem existir atributos cujo intervalo possível de seus valores sejam diferentes. Ou seja, para o atributo "Quantidade de Minúcias Casadas", o intervalo é  $[0, n]$  onde  $n$  é a quantidade de minúcias existentes no template de menor quantidade de minúcias. E para a variável que se refere à fração de casamento o intervalo é de  $[0, 1]$ , onde 1 indica que todas as minúcias da área de intersecção foram pareadas. Dessa forma, a quantidade de minúcias passa a ter peso muito maior que as demais variáveis que tem o seu intervalo de domínio fixo.

Nota-se, entretanto, que utilizando apenas o escopo local de casamento de minúcias e ainda com a deficiência de otimização para tornar o método mais confiável, o método proposto já apresenta resultados satisfatórios, contando com apenas 4,22%, em média, de erro.

Tabela 6.2: Casamento local e global utilizando pesos unitários.

	DB06	BioPass	DB04-2	DB04-1	DB02	DB00	Média
EER	0,23%	1,73%	3,71%	4,60%	1,64%	3,32%	2,54%

Ao adicionar a abordagem de casamento global, que utiliza a transformada fornecida pelo método de casamento local para encontrar o melhor alinhamento entre os conjuntos de minúcias, observa-se uma melhora considerável de aproximadamente 40%, como pode

ser visto na Tabela 6.2. Isso acontece pois o casamento de minúcias utilizando o escopo local é prejudicado por imagens de baixa qualidade que apresentam grande quantidade de minúcias ruidosas. Dessa forma, devido à resistência da abordagem de casamento global a esse tipo de falha, há a diminuição da taxa de erros. Assim, pode-se concluir que aproximadamente 40% dos erros do casamento local podem ser corrigidos quando este método se une ao casamento global.

## 6.2 Otimização dos Pesos dos Atributos de Pontuação

Nesta seção são exibidas as etapas de treinamento utilizando o Algoritmo de Otimização por Nuvens de Partículas e Algoritmos Genéticos. O objetivo da otimização é encontrar o conjunto de parâmetros que, associados às variáveis que compõem o mecanismo de pontuação de similaridade do método, encontre a menor taxa de erro.

O primeiro algoritmo bioinspirado executado foi o PSO e seus resultados estão descritos na Tabela 6.3. Nela, as linhas representam os bancos de dados usados para treinamento e as colunas indicam a aplicação dos dados obtidos a partir do treinamento em outros bancos de dados. A última coluna mostra a média das taxas de erros obtidas em todas as instâncias testadas.

Esta etapa de treinamento durou em média 370 minutos com 8 *threads*, o que tornou a computação proporcionalmente mais rápida.

Por maior que fossem as bases de dados nas quais o algoritmo foi treinado, este algoritmo permaneceu com sua característica de rápida convergência para resultados. Como esta execução não contou com bibliotecas de terceiros, foi possível observar que os resultados convergiam nas primeiras iterações de execução do algoritmo, o que ocorreu na maioria das vezes antes da oitava época.

Na Tabela 6.4, da mesma forma que na tabela anterior, está descrito o processo de treinamento do algoritmo utilizando Algoritmos Genéticos.

Tabela 6.3: Treinamento utilizando Otimização por Nuvens de Partículas

PSO	DB06	BioPass	DB04-2	DB04-1	DB02	DB00	Média
DB06	0,18%	0,98%	3,46%	4,49%	0,67%	3,64%	2,24%
BioPass	0,28%	0,75%	3,78%	4,67%	1,07%	3,54%	2,35%
DB04-2	0,18%	1,30%	3,28%	4,49%	0,75%	3,28%	2,21%
DB04-1	0,21%	0,95%	3,42%	4,03%	0,85%	3,28%	2,12%
DB02	0,19%	1,43%	3,85%	4,74%	1,10%	3,45%	2,46%
DB00	0,28%	1,66%	3,88%	4,85%	1,81%	2,63%	2,52%

Este, por sua vez, teve o processo de treinamento mais demorado pois a biblioteca utilizada não permitia que o trabalho fosse otimizado com o uso de *threads*. O treinamento completo chegou a ser executado por 1.453 minutos, pouco mais de 24h.

Como os algoritmos genéticos têm a habilidade de percorrer as dimensões de busca com mais eficiência devido a possibilidade de mutação e criação de novos cromossomos ao decorrer das gerações, pode-se notar que foi possível encontrar melhores resultados nos bancos de dados treinados, como pode-se observar na diagonal principal da Tabela 6.4.

Tabela 6.4: Treinamento utilizando Algoritmos Genéticos

GA	DB06	BioPass	DB04-2	DB04-1	DB02	DB00	Média
DB06	0,16%	1,23%	3,78%	4,62%	0,61%	4,06%	2,41%
BioPass	0,24%	0,71%	3,78%	4,89%	1,07%	3,89%	2,43%
DB04-2	0,30%	1,72%	3,42%	4,85%	1,38%	3,14%	2,47%
DB04-1	0,25%	1,66%	3,71%	4,13%	0,89%	3,73%	2,39%
DB02	0,24%	1,45%	4,39%	5,27%	0,25%	4,78%	2,73%
DB00	0,42%	1,97%	4,17%	5,42%	2,39%	2,85%	2,87%

Entretanto, um treinamento cuja função possui um ajuste muito adaptado ao seu banco de treinamento pode gerar o fenômeno de super-ajuste. Dessa forma, esse algo-

ritmo conseguiu resultados melhores no banco de treinamento mas, ao ser testado nas demais bases de dados, não obteve o mesmo desempenho. O que pode ser comprovado com o aumento das médias das taxas de erros.

O treinamento que obteve menor média de erro dentre as bases de dados testada foi o treinamento na base de dados DB04-1 utilizando o método de treinamento PSO. Esse desempenho foi, aproximadamente, 17% melhor que o método proposto utilizando pesos unitários para a formação da função de pontuação de similaridade e obteve EER 50% menor que o uso do método de casamento local isoladamente.

### 6.2.1 Aplicação do Método Envoltória Convexa

A utilização do método envoltória convexa, no algoritmo proposto, obteve um desempenho 10% melhor do que o uso do método utilizando fração simples proposto por Maltoni et al. (2009) e Medina-Pérez et al. (2011), no caso em que estes apresentaram menor EER alcançado pela metodologia de treinamento.

Tabela 6.5: Método proposto utilizando cálculo de fração sem o cálculo da envoltória convexa.

	DB06	BioPass	DB04-2	DB04-1	DB02	DB00	Média
DB06	0,18%	0,95%	3,56%	4,77%	1,07%	3,59%	2,35%
BioPass	0,29%	0,71%	3,85%	5,56%	0,99%	4,95%	2,72%
DB04-2	0,21%	1,15%	3,54%	4,63%	1,38%	3,14%	2,34%
DB04-1	0,24%	1,70%	3,92%	4,35%	1,56%	3,42%	2,53%
DB02	0,28%	1,23%	4,39%	7,56%	0,58%	5,95%	3,33%
DB00	0,29%	1,66%	3,78%	4,95%	1,89%	2,66%	2,54%

A Tabela 6.5 mostra o desempenho do algoritmo proposto utilizando o método de fração simples de comparação e usando o PSO como método de treinamento. Comparativamente à Tabela 6.3, que utiliza o método da envoltória convexa, a Tabela 6.5 apresenta pior desempenho em todas as instâncias de treinamento. A média geral, que

representa a média de todos treinamentos apresenta ganho de 13% quando aplicado o método da envoltória convexa.

A Tabela 6.6 sumariza a comparação entre o desempenho dos métodos que utilizam a envoltória convexa e o método de fração simples de comparação. A linha que indica Razão de Desempenho informa a razão entre a média alcançada pelo método da envoltória convexa e a fração de casamento simples, representados pelas linhas Envoltória Convexa e Fração de Casamento, respectivamente.

Tabela 6.6: Comparação de desempenho no algoritmo proposto utilizando o método da envoltória convexa

	DB06	BioPass	DB04-2	DB04-1	DB02	DB00	Média
Envoltória Convexa	2,24%	2,35%	2,21%	2,12%	2,46%	2,52%	2,32%
Fração de Casamento	2,35%	2,72%	2,34%	2,53%	3,33%	2,54%	2,64%
Razão de Desempenho	0,95	0,86	0,94	0,83	0,73	0,99	0,87

Observa-se que foi possível obter ganhos de até 27%, no caso do banco DB02, alcançando uma média de 13% de melhoria de desempenho.

O impacto desta função pode ser ainda maior em algoritmos que utilizam apenas esse critério para definir o grau de similaridade entre duas amostras. Neste trabalho, a variável que mede a similaridade utilizando a fração de casamento é parte de uma função maior que contém outras variáveis.

A Tabela 6.7 sumariza o desenvolvimento do método proposto apresentando o desempenho do algoritmo utilizando as metodologias propostas de modo discriminado.

Nas linhas que indicam que houve treinamento, foi selecionado o treinamento cujo conjunto de testes apresentou menor média de EER.

Tabela 6.7: Desempenho discriminado do algoritmo utilizando as metodologias propostas.

EER	DB06	BioPass	DB04-2	DB04-1	DB02	DB00	Média
Local	1,05%	3,17%	5,92%	6,60%	4,84%	3,71%	4,22%
Local + Global	0,23%	1,73%	3,71%	4,60%	1,64%	3,32%	2,54%
Local + Global + Otimização	0,21%	1,15%	3,54%	4,63%	1,38%	3,14%	2,34%
Local + Global + Envoltória Convexa + Otimização	0,21%	0,95%	3,42%	4,03%	0,85%	3,28%	2,12%

### 6.3 Comparação com Outros Trabalhos

Os algoritmos de casamento de impressões digitais *Minutia Cylinder Code* (MCC), *Source AFIS* e M3GL, que têm suas ferramentas disponibilizadas para fins de pesquisa, foram executados nas bases de dados utilizadas neste trabalho com a finalidade de medir o seu desempenho e compará-los com o método proposto nesta pesquisa. A Tabela 6.8 mostra o desenho daqueles nas bases de dados apresentadas.

Tabela 6.8: Comparação de EER's entre algoritmos de casamento de impressões digitais.

Algoritmo/EER	DB06	BioPass	DB04-2	DB04-1	DB02	DB00	Média
M3GL	1,52%	7,15%	17,87%	19,98%	19,62%	10,32%	12,74%
Source AFIS	0,34%	1,71%	4,75%	5,35%	1,24%	6,45%	3,31%
MCC	0,17%	1,15%	3,93%	3,19%	0,57%	2,99%	2,00%
Método Proposto	0,21%	0,95%	3,42%	4,03%	0,85%	3,28%	2,12%

Observou-se que o método proposto e o algoritmo MCC obtiveram desempenho superior aos demais algoritmos avaliados. Assim, analisando-os separadamente, a Tabela 6.9 dispõe o desempenho do algoritmo MCC e o método proposto de forma comparativa. As células que estão com fonte de cor esverdeada indicam que o respectivo algoritmo

obteve melhor desempenho do que o outro no conjunto indicado pelo título da coluna. A terceira linha indica a razão de desempenho do algoritmo do MCC em relação ao método proposto.

Tabela 6.9: Comparação entre os algoritmos MCC e o método proposto

Algoritmo/EER	DB06	BioPass	DB04-2	DB04-1	DB02	DB00	Média
MCC	<b>0,17%</b>	1,15%	3,93%	<b>3,19%</b>	<b>0,57%</b>	<b>2,99%</b>	<b>2,00%</b>
Método Proposto	0,21%	<b>0,95%</b>	<b>3,42%</b>	4,03%	0,85%	3,28%	2,12%
Dempenho	17%	-21%	-15%	21%	33%	9%	6%

Observa-se que o desempenho de ambos foi similar. Apesar de o algoritmo MCC obter melhor desempenho na maioria das bases de dados testadas, em média, aquele algoritmo apresentou desempenho apenas 6% superior ao método proposto.

A Tabela 6.10 mostra o desempenho do método proposto, fonte de cor esverdeada, em relação aos demais candidatos de todas as versões do projeto FVC. Observa-se que o método proposto alcançaria a primeira colocação no projeto FVC 2000, que teve um total de 11 participantes; oitava colocação no projeto FVC 2002 que teve um total de 31 participantes; sexta colocação em ambas modalidades do projeto FVC 2004-1, onde 41 participaram da modalidade do banco DB1 e 42 participaram da modalidade do banco DB2; e na décima colocação no projeto FVC 2006 que teve um total de 42 participantes classificados.

Atualmente o algoritmo está na vigésima posição do projeto FVC-OnGoing com taxa de erro de 0,67%. Nesta competição, que está em vigor desde a última versão do projeto, o banco de dados no qual o algoritmo será executado não é conhecido e o método proposto é o único algoritmo classificado que é de origem brasileira. A versão atual do projeto FVC exibe apenas a posição dos 30 melhores algoritmos classificados, pela grande quantidade de submissões que o projeto recebe.

As tabelas completas retiradas dos sites dos projetos estão disponíveis no Apêndice

B deste trabalho.

O método proposto tem o tempo de execução médio de 167 milissegundos por casamento de impressões digitais. Uma vantagem deste tipo de algoritmo é a capacidade de executar os casamentos de forma que os processos não apresentem dependência entre si, o que permite a paralelização do problema.

No capítulo seguinte serão feitas as considerações finais a respeito do trabalho apresentado e os trabalhos futuros.

Tabela 6.10: Desempenho do método proposto, de cor verde, em relação aos demais candidatos em todas as versões do projeto FVC.

Pos/EER	FVC-OnGoing	DB06	DB04-2	DB04-1	DB02	DB00
1	0,19%	0,02%	1,58%	1,97%	0,10%	<b>3,28%</b>
2	0,23%	0,03%	2,59%	2,72%	0,24%	3,64%
3	0,23%	0,10%	2,79%	3,38%	0,25%	4,01%
4	0,25%	0,10%	3,17%	3,62%	0,52%	5,36%
5	0,25%	0,12%	3,23%	3,91%	0,61%	8,29%
6	0,29%	0,12%	<b>3,42%</b>	<b>4,03%</b>	0,63%	11,94%
7	0,32%	0,14%	3,49%	4,10%	0,67%	12,20%
8	0,35%	0,14%	3,56%	4,18%	<b>0,85%</b>	16,32%
9	0,38%	0,19%	3,62%	4,37%	0,98%	17,73%
10	0,39%	<b>0,21%</b>	3,75%	5,54%	1,17%	22,63%
11	0,41%	0,24%	4,01%	5,64%	1,46%	23,18%
12	0,44%	0,25%	4,39%	6,13%	1,63%	47,43%
13	0,48%	0,27%	4,67%	6,49%	1,85%	-
14	0,51%	0,29%	4,83%	7,18%	1,91%	-
15	0,52%	0,37%	4,99%	7,47%	2,15%	-
16	0,54%	0,47%	5,15%	7,61%	2,36%	-
17	0,58%	0,49%	5,48%	7,65%	2,57%	-
18	0,60%	0,51%	5,50%	7,68%	2,70%	-
19	0,65%	0,49%	5,54%	8,31%	2,72%	-
20	<b>0,67%</b>	0,66%	5,68%	8,41%	3,02%	-

## Capítulo 7

# Considerações Finais

Este trabalho apresentou uma nova abordagem de casamento de impressões digitais baseado em minúcias, etapa importante na identificação de indivíduos utilizando este traço biométrico, seguindo as normas do padrão ISO/IEC 19794-2:2005. O método proposto apresenta uma etapa de casamento de abordagem local, abordagem de casamento global e otimização de parâmetros utilizando dois algoritmos bioinspirados, os algoritmos genéticos e o método de otimização por nuvens de partículas. A validação deste método é realizada calculando o EER nos bancos de impressões digitais disponibilizados pelas versões anteriores do projeto FVC.

Para os experimentos foram utilizadas 4800 imagens de impressões digitais que tiveram seus atributos extraídos pelo mesmo algoritmo de extração de características. Essas instâncias pertencem a bancos de dados das versões do projeto FVC do ano 2000, 2002, 2004, 2006 e um banco de imagens do projeto Biopass-UFPB.

O processo de otimização de características utilizando algoritmos bioinspirados otimizou um total de 14 atributos pertencentes à etapa de casamento local, além de pesos atribuídos às etapas local e global de casamento a fim de unir as duas abordagens, tornando o algoritmo mais robusto. A função objetivo desse processo é a diminuição da taxa de erro, EER, de forma que cada conjunto de parâmetro aplicado através

dos otimizadores realizam o cálculo do EER, executa-se o algoritmo de casamento de impressões digitais por todo o banco de dados e retorna o valor obtido. Dessa forma, quanto menor a taxa de erro, melhor é o resultado obtido.

O conjunto de treinamento e teste seguiu a metodologia de validação cruzada, onde todos os bancos de dados são treinados uma vez, e todos os demais são testados a cada treinamento. Apesar da pequena quantidade de atributos a serem otimizados, os treinamentos chegaram a durar mais de 24h ininterruptas, pois a cada verificação de taxa de erro era necessário recalcular essa taxa em todo o banco de dados com o objetivo de obter a taxa de erro para o dado conjunto de parâmetros.

O estudo apresentado mostrou, desde antes da otimização de parâmetros, taxas de acerto de aproximadamente 98%. Os resultados obtidos após o treinamento apresentaram melhoria de 50% sobre os resultados iniciais apresentados neste trabalho.

Observou-se que a aplicação de ambos algoritmos bioinspirados são eficazes para a otimização, entretanto, os algoritmos genéticos apresentaram uma tendência ao superajustamento, obtendo melhores resultados nos bancos treinados, porém os demais bancos testados não obtiveram resultados tão bons. Entretanto o algoritmo de otimização por nuvens de partículas se mostrou mais estável chegando a alcançar uma taxa de acerto de 99,8% no banco de dados do projeto FVC 2006 e uma média de erro de 2,12% nos bancos de dados testados.

Ao comparar o método proposto com o algoritmo considerado estado da arte, observou-se uma diferença de desempenho de aproximadamente 6%. Onde o algoritmo MCC, considerado estado da arte, obteve melhor desempenho em 4 dos 6 bancos apresentados, obtendo um desempenho médio de 2% de erro contra 2,12% do método proposto.

Ao analisar a possível colocação do método proposto nas versões anteriores do projeto FVC, que realiza o *benchmarking* dos algoritmos de casamento de impressões digitais submetidos àquela competição, obteve-se a primeira colocação, no FVC 2000;

oitava colocação, no FVC 2002; sexta colocação, nos dois bancos testados da versão 2004 do mesmo projeto e a décima colocação na versão 2006 do projeto FVC.

O algoritmo desenvolvido foi construído baseado nas diversas metodologias disponíveis na literatura e através de pesquisa científica. Na etapa de casamento local, verificou-se que um cálculo, utilizado por diversos autores, apresentava uma falha. Áreas das amostras que não faziam parte da intersecção das impressões digitais estavam sendo consideradas na etapa de cálculo de similaridade entre elas. O problema foi corrigido através da aplicação do método da envoltória convexa para identificar apenas as informações presentes na área de intersecção, o que resultou em uma melhoria de aproximadamente 13% comparado ao método convencional proposto por Maltoni et al. (2009) e outros autores.

Atualmente o método proposto está na vigésima colocação na versão FVC-*OnGoing*, que está em vigor desde a última versão, com um desempenho de 0,67% de EER. Atualmente não existem algoritmos de origem brasileira com resultados apresentados nesta competição além do método proposto.

Em suma, diante do que foi exposto, pode-se considerar que este trabalho atingiu os objetivos esperados. Não obstante, embora a pesquisa científica utilizando este traço biométrico já tenha avançado por longos anos, o método proposto é capaz de competir com os algoritmos que disputam o mercado, de acordo com os resultados apresentados em uma competição com algoritmos reais. Este trabalho apresenta um grande avanço para a biometria em âmbito nacional, tendo em vista que não há algoritmos de origem brasileira publicados nesta competição.

Para trabalhos futuros, uma nova categoria de casamento de impressões digitais pode ser explorada, visto que este trabalho apresentou um método de casamento de impressões digitais baseado em minúcias seguindo as normas da ISO/IEC 19794-2:2005, a etapa de casamento seguindo estas normas dispõe de apenas um pequeno conjunto de informações que pertencem às minúcias. Portanto, existem outras informações que

podem ser extraídas das impressões digitais e tornar este método ainda mais confiável.

# Referências Bibliográficas

- Anderson, C. e Franks, N. R. (2001). Teams in animal societies. *Behavioral Ecology*, 12(5):534–540.
- Andrezza, I. L. P. et al. (2015). Análise de técnicas de normalização aplicadas ao reconhecimento facial.
- Bäck, T. e Schwefel, H.-P. (1993). An overview of evolutionary algorithms for parameter optimization. *Evolutionary computation*, 1(1):1–23.
- Barber, C. B., Dobkin, D. P., e Huhdanpaa, H. (1996). The quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software (TOMS)*, 22(4):469–483.
- Behroozmand, R. e Almasganj, F. (2007). Optimal selection of wavelet-packet-based features using genetic algorithm in pathological assessment of patients' speech signal with unilateral vocal fold paralysis. *Computers in Biology and Medicine*, 37(4):474–485.
- BioLab (2006). <https://biolab.csr.unibo.it/fvcongoing> . acesso: julho de 2016.
- Bolle, R. (2004). *Guide to Biometrics*. Springer Science & Business Media.
- Bolle, R. M., Connell, J. H., e Ratha, N. K. (2002). Biometric perils and patches. *Pattern Recognition*, 35(12):2727–2738.

- Brahnam, S., Nanni, L., e Lumini, A. (2013). A secure multimatcher system for fingerprint verification. *Proceedings for the Northeast Region Decision Sciences Institute*.
- Cappelli, R., Ferrara, M., e Maltoni, D. (2010a). Minutia cylinder-code: A new representation and matching technique for fingerprint recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(12):2128–2141.
- Cappelli, R., Ferrara, M., e Maltoni, D. (2012). Minutiae-based fingerprint matching. Em *Cross Disciplinary Biometric Systems*, pgs. 117–150. Springer.
- Cappelli, R., Ferrara, M., Maltoni, D., e Tistarelli, M. (2010b). Mcc: A baseline algorithm for fingerprint verification in fvc-ongoing. Em *Control Automation Robotics & Vision (ICARCV), 2010 11th International Conference on*, pgs. 19–23. IEEE.
- Cappelli, R., Maltoni, D., e Ferrara, M. (2011). Fingerprint indexing based on minutia cylinder-code. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(5):1051–1057.
- Chikkerur, S. S. (2005). *Online fingerprint verification system*. PhD thesis, Citeseer.
- Clarke, R. (1994). Human identification in information systems: Management challenges and public policy issues. *Information Technology & People*, 7(4):6–37.
- Coelho Filho, O. P., Martinhon, C. A., e Cabral, L. d. A. F. (2013). Uma abordagem melhorada do algoritmo de otimização por enxame de partículas para o problema de clusterização de dados. *Simposio Brasileiro de Pesquisa Operacional*, pgs. 2135–2147.
- Committees, I. T. (2005). Iso/iec 19794-2:2005, information technology—biometric data interchange formats—part 2: Finger minutiae data. ISO 19794-2:2005, International Organization for Standardization, Geneva, Switzerland.
- Costa, L. R., Obelheiro, R. R., e Fraga, J. S. (2006). Introdução á biometria. *Livro*

- texto dos Minicursos do VI Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais (SBSeg2006)*. SBC: Porto Alegre, 1:103–151.
- Dalila, C., Imane, H., e Amine, N.-A. (2015). Multimodal score-level fusion using hybrid ga-pso for multibiometric system. *Informatica*, 39(2).
- Darwin, C. (1872). *The origin of species*.
- Dawid, H. (2011). *Adaptive learning by genetic algorithms: Analytical results and applications to economic models*. Springer Science & Business Media.
- De Castro, L. N. e Von Zuben, F. J. (2005). *Recent developments in biologically inspired computing*. Igi Global.
- Dietterich, T. G. (2000). Ensemble methods in machine learning. Em *International workshop on multiple classifier systems*, pgs. 1–15. Springer.
- Duraisamy, S. P. e Kayalvizhi, R. (2010). A new multilevel thresholding method using swarm intelligence algorithm for image segmentation. *Journal of Intelligent Learning Systems and Applications*, 2(03):126.
- Eberhart, R. C. e Kennedy, J. (1995). A new optimizer using particle swarm theory. Em *Proceedings of the sixth international symposium on micro machine and human science*, volume 1, pgs. 39–43. New York, NY.
- Ferrara, M., Maltoni, D., e Cappelli, R. (2012). Noninvertible minutia cylinder-code representation. *Information Forensics and Security, IEEE Transactions on*, 7(6):1727–1737.
- França, F. O. (2005). *Algoritmos bio-inspirados aplicados à otimização dinâmica*. PhD thesis, UNIVERSIDADE ESTADUAL DE CAMPINAS.
- Golberg, D. E. (1989). Genetic algorithms in search, optimization, and machine learning. *Addion wesley*, 1989:102.

- Hansen, L. K. e Salamon, P. (1990). Neural network ensembles. *IEEE transactions on pattern analysis and machine intelligence*, 12:993–1001.
- Holland, J. H. (1962). Outline for a logical theory of adaptive systems. *Journal of the ACM (JACM)*, 9(3):297–314.
- Holland, J. H. (1975). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. U Michigan Press.
- Jain, A., Flynn, P., e Ross, A. A. (2007). *Handbook of biometrics*. Springer Science & Business Media.
- Jain, A. K., Ross, A., e Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on circuits and systems for video technology*, 14(1):4–20.
- Kirkpatrick, D. G. e Seidel, R. (1986). The ultimate planar convex hull algorithm? *SIAM journal on computing*, 15(1):287–299.
- Kulp, A. e Braskamp, H. (2012). Selecting identification technology. <http://nedap-securitymanagement.com/en/technology-nedapmenu-90/identificationtechnology-nedapmenu-61.html>. Online; Acessado em Agosto de 2012.
- Lee, H. C., Ramotowski, R., e Gaensslen, R. (2001). *Advances in fingerprint technology*. CRC press.
- Luque-Baena, R. M., Elizondo, D., López-Rubio, E., Palomo, E. J., e Watson, T. (2013). Assessment of geometric features for individual identification and verification in biometric hand systems. *Expert systems with applications*, 40(9):3580–3594.
- Maltoni, D., Maio, D., Jain, A., e Prabhakar, S. (2009). *Handbook of fingerprint recognition*. Springer Science & Business Media.

- McCulloch, W. S. e Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133.
- McLachlan, G. J., Bean, R., e Peel, D. (2002). A mixture model-based approach to the clustering of microarray expression data. *Bioinformatics*, 18(3):413–422.
- Medina-Pérez, M. A., García-Borroto, M., Gutiérrez-Rodríguez, A. E., e Altamirano-Robles, L. (2011). Robust fingerprint verification using m-triplets. Em *Hand-Based Biometrics (ICHB), 2011 International Conference on*, pgs. 1–5. IEEE.
- Medina-Pérez, M. A., Moreno, A. M., Ballester, M. Á. F., García-Borroto, M., Loyola-González, O., e Altamirano-Robles, L. (2016). Latent fingerprint identification using deformable minutiae clustering. *Neurocomputing*, 175:851–865.
- Meffert, K., Rotstan, N., Knowles, C., e Sangiorgi, U. (2012). Jgap-java genetic algorithms and genetic programming package. URL: <http://jgap.sf.net>.
- Mitchell, M. (1995). Genetic algorithms: An overview. *Complexity*, 1(1):31–39.
- Mordini, E. (2005). Biometric identification technology ethics. *BITE Policy Paper*, pg. 64.
- Neurotech (2014). <http://www.neurotechnology.com/> . acesso em julho de 2016.
- O’Hara, C. E. e O’Hara, G. L. (1956). *Fundamentals of criminal investigation*. Thomas Springfield, IL.
- o’Rourke, J. (1998). *Computational geometry in C*. Cambridge university press.
- Paton, R. (1994). *Computing with biological metaphors*. Chapman & Hall.
- Patrick, E. A. (1972). *Fundamentals of pattern recognition*. Prentice Hall.

- Queiroga, E. V. (2016). Meta-heurísticas de otimização contínua para clusterização de dados e segmentação de imagens. *Programa de Pós Graduação em Informática - UFPB*.
- Rahman, M. A., Anwar, F., e Azad, M. S. (2008). A simple and effective technique for human verification with hand geometry. Em *Computer and Communication Engineering, 2008. ICCCE 2008. International Conference on*, pgs. 1177–1180. IEEE.
- Raviraj (2015). Raviraj technologies inc. india. web site.
- Rosenblatt, F. (1957). *The perceptron, a perceiving and recognizing automaton Project Para*. Cornell Aeronautical Laboratory.
- Ross, A. A., Nandakumar, K., e Jain, A. (2006). *Handbook of multibiometrics*, volume 6. Springer Science & Business Media.
- Silva, A. G. A. (2015). Classificação e verificação multinométrica por geometria da mão e impressão palmar com otimização por algoritmos genéticos. *Programa de Pós-Graduação em Informática, UFPB*.
- Sun, Z., Bebis, G., e Miller, R. (2004). Object detection using feature subset selection. *Pattern recognition*, 37(11):2165–2176.
- Thomas Bäck, David B. Fogel, Z. M. (1997). *Handbook of Evolutionary Computation (Computational Intelligence Library)*. Computational Intelligence Library. Published in cooperation with the Institute of Physics, 1st edition.
- Važan, R. (2009). Sourceafis project.
- Weisstein, E. W. (2016). Convex hull.
- Whitley, L. D. (1993). *Foundations of genetic algorithms 2*, volume 2. Morgan Kaufmann.

Zhang, D. D. (2004). *Palmprint authentication*. Springer Science & Business Media.

# Apêndice A

## Templates ISO/IEC 19794-2:2005

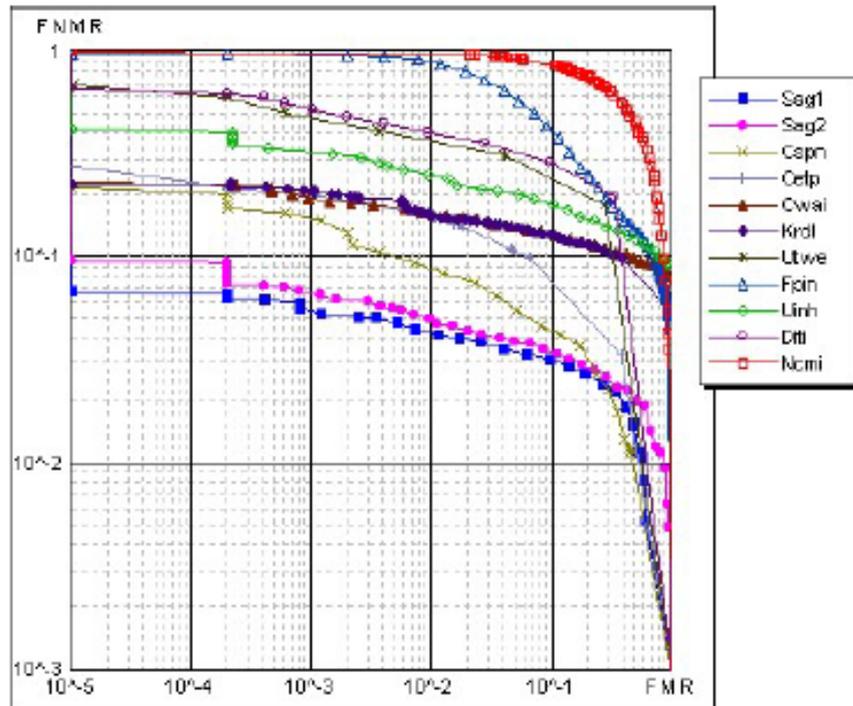
		Field	Size	Valid Values and Notes	
Finger Minutiae Record	Record Header	...	14 bytes		
		Image Horizontal Size	2 bytes	in pixels	
		Image Vertical Size	2 bytes	in pixels	
		Horizontal Resolution	2 bytes	in pixels per cm	
		Vertical Resolution	2 bytes	in pixels per cm	
		Number of Finger Views $n_F$	1 byte	0 to 255	
		...	1 byte		
	Single Finger Record ( $n_F$ instances)	Finger Header	Finger Position	1 byte	0 to 10 (See Table 2)
			View Number	4 bits	0 to 15
			Impression Type	4 bits	0 to 8 (See Table 3)
			Finger Quality	1 byte	1 to 100 (0=quality not reported)
			Number of Minutiae $n$	1 byte	0 to 255
		Finger Minutia Record ( $n$ instances)	Type	2 bits	{00= OTHER, 01= RIDGE_ENDING, 10= RIDGE_BIFURCATION}
			Position $x$	14 bits	in pixels
			Reserved	2 bits	
			Position $y$	14 bits	in pixels
			Orientation $\theta$	1 byte	0 to 255 (resolution 1.40625 degrees)
			Minutia Quality	1 byte	1 to 100 (0=quality not reported)
			Extended Data Block Length	2 bytes	
		Extended Data (0+ inst.)	Extended Data Area Type Code	2 bytes	
Extended Data Area Length			2 bytes	only present if Extended Data Block Length>0	
Data Section	(prev. field)				

Figura A.1: Estrutura de Formação de Templates ISO/IEC 19794-2:2005 de Impressão Digital

## Apêndice B

# Classificação dos Participantes nos eventos do projeto FVC.

### Results on DB3



Algorithm	EER	EER <sup>±</sup>	REJ <sub>ENROLL</sub>	REJ <sub>MATCH</sub>	Avg Enroll Time	Avg Match Time
<i>Sag1</i>	3.64%	3.64%	0.00%	0.00%	5.70 sec	2.13 sec
<i>Sag2</i>	4.01%	4.01%	0.00%	0.00%	1.94 sec	1.94 sec
<i>Cspn</i>	5.36%	5.33%	0.57%	1.24%	0.35 sec	0.36 sec
<i>Cetp</i>	8.29%	8.29%	0.00%	0.00%	1.49 sec	1.66 sec
<i>Cwai</i>	11.94%	5.90%	12.86%	8.00%	0.46 sec	0.57 sec
<i>Krdl</i>	12.20%	8.03%	6.86%	5.12%	1.48 sec	1.60 sec
<i>Uinh</i>	16.32%	10.14%	10.29%	7.64%	1.28 sec	1.36 sec
<i>Utwe</i>	17.73%	17.73%	0.00%	0.00%	10.44 sec	2.31 sec
<i>Diti</i>	22.63%	22.63%	0.00%	0.00%	2.59 sec	2.67 sec
<i>Fpin</i>	23.18%	23.18%	0.00%	0.00%	2.13 sec	2.19 sec
<i>Ncmi</i>	47.43%	47.44%	0.00%	0.01%	2.25 sec	2.75 sec

Figura B.1: Classificação dos algoritmos do projeto FVC 2000.

## Results on DB1

Algorithm	EER	FMR 100	FMR 1000	Zero FMR	REJ <sub>ENROLL</sub>	REJ <sub>MATCH</sub>	Avg Enroll Time	Avg Match Time
PA15	0.10%	0.11%	0.21%	0.21%	0.00%	0.00%	0.12 sec	2.47 sec
PB27	0.24%	0.18%	0.39%	0.82%	0.00%	0.00%	1.30 sec	1.15 sec
PA27	0.25%	0.21%	0.36%	0.86%	0.00%	0.00%	2.26 sec	1.99 sec
PB05	0.52%	0.86%	0.96%	1.11%	0.00%	0.00%	0.54 sec	0.58 sec
PB35	0.61%	0.54%	1.14%	1.36%	0.00%	0.00%	0.92 sec	0.76 sec
PB15	0.63%	0.61%	0.82%	1.04%	0.00%	0.00%	0.07 sec	0.24 sec
PA35	0.67%	0.61%	1.00%	1.64%	0.00%	0.00%	4.82 sec	1.97 sec
PA08	0.98%	1.00%	2.21%	3.50%	0.00%	0.00%	0.52 sec	0.52 sec
PA45	1.17%	1.21%	2.21%	2.79%	0.00%	0.00%	0.55 sec	0.67 sec
PA13	1.46%	1.61%	2.25%	26.36%	0.00%	0.04%	0.17 sec	0.44 sec
PA26	1.63%	1.79%	2.36%	3.57%	0.00%	0.00%	0.53 sec	0.61 sec
PA34	1.85%	2.18%	3.39%	7.43%	0.00%	0.00%	0.56 sec	0.68 sec
PA12	1.91%	2.46%	3.82%	4.18%	0.00%	0.00%	0.26 sec	0.27 sec
PA19	2.15%	2.75%	5.32%	7.96%	0.00%	0.00%	0.19 sec	0.20 sec
PA24	2.36%	2.82%	4.29%	4.64%	0.00%	0.00%	0.58 sec	0.60 sec
PA02	2.57%	2.68%	3.25%	3.82%	1.29%	0.80%	0.83 sec	1.29 sec
PA14	2.70%	3.29%	5.64%	7.93%	0.00%	0.00%	0.80 sec	1.84 sec
PA29	2.72%	4.46%	7.14%	8.93%	0.00%	0.00%	0.67 sec	0.68 sec
PA32	3.02%	3.75%	5.64%	9.07%	0.29%	0.18%	0.34 sec	0.59 sec
PA07	3.74%	6.86%	13.29%	18.68%	0.00%	0.00%	0.21 sec	0.61 sec
PA31	3.85%	6.11%	10.75%	21.57%	0.00%	0.01%	0.01 sec	3.16 sec
PA10	4.37%	4.71%	5.29%	5.79%	0.00%	0.00%	1.91 sec	1.91 sec
PA28	4.78%	9.43%	17.64%	20.89%	0.00%	0.00%	0.52 sec	0.87 sec
PA21	4.97%	7.21%	15.04%	21.04%	1.00%	0.71%	0.69 sec	0.74 sec
PA18	6.12%	13.75%	13.75%	19.39%	0.43%	0.04%	0.70 sec	0.71 sec
PA20	6.71%	9.39%	12.39%	15.86%	0.00%	0.00%	0.13 sec	0.16 sec
PA42	8.27%	16.96%	23.86%	33.14%	0.86%	0.54%	0.51 sec	0.52 sec
PA16	16.28%	24.00%	28.11%	40.82%	0.00%	0.12%	1.22 sec	1.24 sec
PA22	17.34%	40.39%	54.07%	60.93%	0.00%	0.79%	0.59 sec	0.68 sec
PA25	35.00%	83.96%	93.14%	98.64%	0.57%	0.14%	0.56 sec	0.67 sec
PA03	50.00%	100.00%	100.00%	100.00%	0.00%	100.00%	7.59 sec	5.01 sec

Figura B.2: Classificação dos algoritmos do projeto FVC 2002.

Open Category: Results on DB1

Alg	EER	FMR 100	FMR 1000	Zero FMR	REJENROLL	REJMATCH	Avg Enroll Time	Avg Match Time	Avg Model Size	Max Model Size	Max Enroll Mem	Max Match Mem
P047	1.97%	2.86%	4.36%	8.21%	0.00%	0.00%	1.95s	1.87s	1.4Kb	2.8Kb	5080Kb	5796Kb
P101	2.72%	3.86%	9.25%	13.43%	0.00%	0.00%	0.10s	3.19s	31.5Kb	31.5Kb	3204Kb	5464Kb
P097	3.38%	5.54%	9.75%	12.93%	0.00%	0.00%	0.74s	0.75s	17.8Kb	28.1Kb	5564Kb	5780Kb
P009	3.62%	5.54%	11.68%	15.39%	0.00%	0.00%	0.25s	0.23s	1.1Kb	2.9Kb	2828Kb	2860Kb
P049	3.91%	7.11%	12.43%	17.96%	0.00%	0.00%	0.32s	0.47s	0.5Kb	0.9Kb	2472Kb	2496Kb
P004	4.10%	5.64%	8.61%	11.68%	0.00%	0.00%	0.77s	0.75s	1.9Kb	3.2Kb	7012Kb	7032Kb
P103	4.18%	6.36%	11.25%	17.25%	0.00%	0.00%	0.16s	0.16s	1.0Kb	1.0Kb	3588Kb	3684Kb
P071	4.37%	5.43%	9.46%	21.64%	0.00%	0.00%	0.39s	0.77s	17.5Kb	31.4Kb	5872Kb	9800Kb
P026	5.54%	9.25%	19.11%	24.89%	0.00%	0.94%	2.60s	3.56s	8.4Kb	18.4Kb	21300Kb	36392Kb
P075	5.64%	63.43%	64.75%	80.50%	0.00%	0.00%	0.41s	0.44s	0.4Kb	0.9Kb	2348Kb	2352Kb
P108	6.13%	10.89%	15.82%	18.36%	0.00%	0.00%	0.35s	0.35s	1.7Kb	1.7Kb	2840Kb	2848Kb
P083	6.49%	13.11%	19.68%	26.43%	0.00%	66.86%	0.37s	0.37s	0.6Kb	0.6Kb	3380Kb	3392Kb
P039	7.18%	12.71%	16.68%	18.00%	0.00%	0.00%	1.16s	1.32s	3.4Kb	3.9Kb	3976Kb	4084Kb
P048	7.47%	14.14%	18.71%	22.64%	0.00%	0.00%	0.44s	0.40s	0.5Kb	0.7Kb	5888Kb	5920Kb
P041	7.61%	14.43%	20.64%	21.61%	0.00%	0.00%	0.17s	0.20s	9.1Kb	9.1Kb	1032Kb	1068Kb
P113	7.65%	11.68%	15.54%	19.61%	0.00%	0.00%	0.67s	0.69s	53.3Kb	93.2Kb	11936Kb	12260Kb
P068	7.68%	13.86%	21.68%	24.96%	0.00%	0.00%	0.63s	0.65s	7.9Kb	7.9Kb	4468Kb	4456Kb
P111	8.31%	16.82%	28.82%	34.86%	0.00%	0.00%	0.64s	0.64s	1.6Kb	2.1Kb	2032Kb	2036Kb
P052	8.41%	14.82%	21.36%	33.61%	0.14%	0.09%	0.28s	0.26s	3.2Kb	6.5Kb	4380Kb	4268Kb
P016	8.87%	14.64%	20.04%	24.82%	0.00%	0.00%	0.16s	0.35s	14.3Kb	24.2Kb	2396Kb	3712Kb
P120	8.94%	19.04%	26.07%	35.46%	0.00%	0.00%	1.90s	2.00s	0.5Kb	1.1Kb	12144Kb	12492Kb
P099	9.55%	19.07%	30.21%	35.86%	0.00%	0.00%	0.46s	0.44s	0.5Kb	0.5Kb	6488Kb	5920Kb
P027	9.78%	19.79%	30.54%	55.32%	0.00%	0.00%	1.05s	3.26s	4.9Kb	5.0Kb	5156Kb	7568Kb
P050	10.06%	16.43%	22.79%	25.50%	0.00%	0.00%	0.67s	0.66s	1.3Kb	1.3Kb	5324Kb	5332Kb
P011	10.39%	16.36%	22.04%	26.39%	0.00%	0.00%	0.70s	0.69s	1.6Kb	1.6Kb	2780Kb	2784Kb
P112	10.57%	24.04%	40.82%	100.00%	0.14%	0.09%	0.35s	0.36s	1.3Kb	1.3Kb	5276Kb	7720Kb
P072	10.81%	23.43%	37.54%	41.21%	0.00%	0.00%	0.12s	0.12s	2.8Kb	3.5Kb	2104Kb	2108Kb
P002	10.85%	23.57%	34.36%	35.71%	0.00%	0.00%	2.13s	2.84s	42.9Kb	44.1Kb	39968Kb	42180Kb
P105	11.25%	26.29%	37.11%	47.75%	0.00%	0.00%	0.37s	0.36s	0.8Kb	1.7Kb	1932Kb	1936Kb
P051	11.93%	33.68%	73.64%	91.86%	0.00%	0.03%	0.14s	0.18s	5.3Kb	5.3Kb	4144Kb	6740Kb
P067	12.48%	24.39%	31.64%	45.14%	0.00%	0.18%	0.07s	0.08s	3.0Kb	3.2Kb	1288Kb	1308Kb
P080	13.53%	26.89%	29.07%	34.00%	0.00%	73.48%	0.62s	0.61s	0.5Kb	0.5Kb	3048Kb	3052Kb
P078	13.65%	35.11%	46.54%	49.82%	0.00%	0.00%	0.19s	0.18s	0.4Kb	0.8Kb	4976Kb	4988Kb
P087	14.03%	31.21%	42.54%	50.64%	0.00%	0.00%	0.28s	0.28s	0.1Kb	0.3Kb	1920Kb	1988Kb
P093	15.30%	30.43%	37.29%	49.21%	0.00%	0.00%	0.08s	0.09s	0.9Kb	2.0Kb	1176Kb	1188Kb
P006	19.36%	45.86%	62.68%	74.36%	0.00%	1.37%	0.34s	0.32s	64.5Kb	64.5Kb	2200Kb	3096Kb
P118	25.69%	49.96%	63.46%	76.11%	0.00%	7.57%	2.90s	3.85s	2400.0Kb	2400.0Kb	4336Kb	7376Kb
P109	35.30%	90.71%	98.46%	99.86%	0.57%	0.19%	1.47s	1.24s	1152.9Kb	2143.0Kb	7316Kb	8116Kb
P106	35.68%	83.89%	92.86%	94.18%	0.00%	0.00%	0.26s	0.34s	2.3Kb	2.3Kb	6372Kb	6404Kb
P079	41.94%	60.25%	66.50%	68.46%	0.00%	0.00%	1.92s	2.78s	512.0Kb	512.0Kb	11968Kb	13512Kb
P104	50.00%	100.00%	100.00%	100.00%	0.00%	100.00%	7.93s	7.99s	0.4Kb	0.9Kb	15980Kb	16100Kb

Figura B.3: Classificação dos algoritmos do projeto FVC 2004 banco de dados DB1.

## Open Category: Results on DB2

Alg	EER	FMR 100	FMR 1000	Zero FMR	REJENROLL	REJMATCH	Avg Enroll Time	Avg Match Time	Avg Model Size	Max Model Size	Max Enroll Mem	Max Match Mem
<a href="#">P039</a>	1.58%	2.18%	5.79%	8.18%	0.00%	0.00%	0.72s	0.83s	2.3Kb	3.0Kb	3268Kb	3384Kb
<a href="#">P071</a>	2.59%	3.14%	4.96%	6.57%	0.00%	0.00%	0.32s	0.57s	14.5Kb	25.9Kb	5656Kb	8720Kb
<a href="#">P004</a>	2.79%	3.79%	5.04%	6.68%	0.00%	0.00%	0.61s	0.62s	2.0Kb	3.8Kb	5260Kb	5376Kb
<a href="#">P113</a>	3.17%	5.93%	7.00%	7.86%	0.00%	0.00%	0.33s	0.35s	40.9Kb	59.5Kb	5080Kb	5308Kb
<a href="#">P097</a>	3.23%	4.25%	6.93%	12.39%	0.14%	0.09%	0.49s	0.53s	13.1Kb	16.2Kb	3100Kb	4144Kb
<a href="#">P047</a>	3.49%	4.46%	7.64%	8.86%	0.00%	0.00%	2.12s	2.18s	1.1Kb	2.3Kb	3632Kb	4348Kb
<a href="#">P101</a>	3.56%	4.36%	5.36%	5.96%	0.00%	0.00%	0.07s	0.81s	18.5Kb	18.5Kb	2324Kb	6076Kb
<a href="#">P049</a>	3.62%	6.04%	8.75%	13.79%	0.00%	0.00%	0.34s	0.34s	0.4Kb	0.6Kb	1772Kb	1796Kb
<a href="#">P068</a>	3.75%	5.54%	8.71%	12.82%	0.00%	0.00%	0.34s	0.39s	7.9Kb	7.9Kb	3804Kb	3788Kb
<a href="#">P009</a>	4.01%	5.43%	7.39%	12.00%	0.00%	0.00%	0.23s	0.23s	1.0Kb	2.2Kb	2136Kb	2172Kb
<a href="#">P016</a>	4.39%	5.50%	7.50%	8.11%	0.00%	0.00%	0.14s	0.31s	14.0Kb	21.4Kb	2124Kb	3436Kb
<a href="#">P048</a>	4.67%	7.75%	13.46%	17.21%	0.00%	0.00%	0.37s	0.36s	0.5Kb	0.6Kb	4140Kb	4148Kb
<a href="#">P108</a>	4.83%	7.11%	10.68%	12.36%	0.00%	0.00%	0.30s	0.31s	1.7Kb	1.7Kb	2464Kb	2468Kb
<a href="#">P103</a>	4.99%	8.25%	12.46%	18.46%	0.00%	0.00%	0.12s	0.12s	1.0Kb	1.0Kb	3036Kb	3280Kb
<a href="#">P041</a>	5.15%	7.43%	9.11%	10.18%	0.00%	0.00%	0.12s	0.14s	3.5Kb	3.5Kb	852Kb	888Kb
<a href="#">P011</a>	5.48%	7.07%	10.43%	22.68%	0.00%	0.00%	0.37s	0.37s	1.6Kb	1.6Kb	1536Kb	1564Kb
<a href="#">P027</a>	5.50%	9.04%	14.07%	20.29%	0.00%	0.00%	0.72s	2.46s	4.9Kb	5.0Kb	3532Kb	4940Kb
<a href="#">P111</a>	5.54%	8.96%	21.71%	59.14%	0.00%	0.00%	0.34s	0.34s	0.8Kb	1.2Kb	1232Kb	1508Kb
<a href="#">P075</a>	5.68%	60.04%	61.86%	69.86%	0.00%	0.00%	0.25s	0.30s	0.3Kb	0.9Kb	1328Kb	1332Kb
<a href="#">P026</a>	5.75%	10.57%	21.46%	100.00%	0.00%	0.46%	2.24s	3.26s	8.0Kb	17.3Kb	18212Kb	36016Kb
<a href="#">P083</a>	6.00%	11.00%	15.39%	18.71%	0.00%	66.99%	0.40s	0.43s	0.6Kb	0.6Kb	2428Kb	2440Kb
<a href="#">P052</a>	6.25%	8.68%	11.71%	16.29%	0.00%	0.00%	0.23s	0.22s	3.2Kb	11.1Kb	3024Kb	2928Kb
<a href="#">P050</a>	6.35%	9.75%	13.36%	15.54%	0.00%	0.00%	0.51s	0.51s	1.3Kb	1.3Kb	4384Kb	4392Kb
<a href="#">P104</a>	6.66%	10.29%	14.82%	16.86%	0.00%	0.12%	2.84s	2.93s	0.5Kb	1.5Kb	8124Kb	18272Kb
<a href="#">P120</a>	6.80%	10.54%	18.00%	23.50%	0.00%	0.03%	1.01s	1.08s	0.4Kb	1.4Kb	6228Kb	11188Kb
<a href="#">P105</a>	7.45%	13.36%	19.96%	22.61%	0.00%	0.00%	0.36s	0.36s	0.9Kb	1.3Kb	1232Kb	1336Kb
<a href="#">P072</a>	7.47%	12.46%	22.61%	28.00%	0.00%	0.00%	0.09s	0.10s	2.7Kb	3.4Kb	1584Kb	1588Kb
<a href="#">P112</a>	7.73%	16.32%	33.18%	100.00%	0.43%	1.54%	0.26s	0.29s	1.3Kb	1.3Kb	5096Kb	7532Kb
<a href="#">P002</a>	7.89%	14.50%	23.39%	26.46%	0.00%	0.01%	0.96s	1.36s	17.2Kb	18.8Kb	22404Kb	29476Kb
<a href="#">P080</a>	8.03%	15.82%	17.86%	20.61%	0.00%	69.43%	0.33s	0.33s	0.5Kb	0.5Kb	2052Kb	2056Kb
<a href="#">P067</a>	8.65%	14.75%	21.21%	25.96%	0.00%	0.00%	0.04s	0.04s	1.3Kb	1.5Kb	1036Kb	1056Kb
<a href="#">P087</a>	8.68%	15.68%	22.75%	32.68%	0.00%	0.00%	0.22s	0.24s	0.1Kb	0.3Kb	1852Kb	1976Kb
<a href="#">P099</a>	8.87%	13.43%	17.46%	19.18%	1.00%	1.91%	0.47s	0.48s	0.5Kb	0.5Kb	4956Kb	4344Kb
<a href="#">P118</a>	9.11%	24.46%	36.00%	60.68%	0.00%	0.00%	2.23s	2.65s	932.8Kb	932.8Kb	2068Kb	3280Kb
<a href="#">P093</a>	11.77%	23.46%	23.46%	37.75%	0.00%	0.00%	0.07s	0.08s	0.9Kb	2.0Kb	984Kb	996Kb
<a href="#">P051</a>	12.92%	36.68%	68.75%	82.50%	0.00%	0.00%	0.12s	0.16s	5.3Kb	5.3Kb	3964Kb	4004Kb
<a href="#">P006</a>	13.63%	29.96%	48.79%	58.29%	0.00%	0.00%	0.29s	0.29s	64.5Kb	64.5Kb	2012Kb	2912Kb
<a href="#">P078</a>	14.36%	36.36%	40.21%	46.75%	0.00%	0.00%	0.12s	0.11s	0.2Kb	0.5Kb	3188Kb	3204Kb
<a href="#">P106</a>	31.05%	79.11%	87.00%	90.61%	0.00%	1.01%	0.20s	0.26s	2.3Kb	2.3Kb	3404Kb	3416Kb
<a href="#">P109</a>	33.80%	89.79%	97.50%	98.18%	0.00%	0.00%	1.31s	1.25s	928.9Kb	1318.5Kb	5292Kb	5320Kb
<a href="#">P079</a>	36.02%	56.93%	63.46%	66.43%	0.00%	0.00%	1.77s	2.70s	512.0Kb	512.0Kb	11784Kb	13332Kb

Figura B.4: Classificação dos algoritmos do projeto FVC 2004 banco de dados DB2.

**Fingerprint Matching (ISO)**

Published on	Benchmark	Participant	Type	Algorithm	Version	EER ▲	FMR <sub>1000</sub>	FMR <sub>10000</sub>	Show details
12/06/2014	FMISO-STD-1.0	Neurotechnology	Company	MM_FMISO	5.1	0,194%	0,328%	0,776%	
15/05/2011	FMISO-STD-1.0	AA Technology Ltd.	Company	EMB9200	2.41	0,234%	0,292%	0,444%	
24/03/2011	FMISO-STD-1.0	UnionCommunity	Company	Triple_M_ISO	1.2	0,234%	0,361%	0,620%	
22/09/2015	FMISO-STD-1.0	Xiamen Toyonway Intellectual Technology Co. Ltd, China	Company	TW2F_ISO	0.2	0,252%	0,314%	0,556%	
15/12/2010	FMISO-STD-1.0	Suprema, Inc.	Company	SFCore	1.0	0,258%	0,346%	0,639%	
09/03/2014	FMISO-STD-1.0	Tiger IT Bangladesh	Company	TigerAFIS	v1.2-ISO/MINEX	0,296%	0,422%	0,837%	
12/10/2009	FMISO-STD-1.0	Tiger IT Bangladesh	Company	Tiger ISO	0.1	0,317%	0,447%	0,866%	

Figura B.5: Classificação dos algoritmos do projeto FVC 2006, classificação do primeiro ao sétimo.

31/12/2015	FMISO-STD-1.0	BKIC Laboratory - Hanoi University of Science and Technology	Academic Research Group	BKAFIS	0.4	0,346%	0,491%	0,696%	
14/05/2011	FMISO-STD-1.0	Institute of Automation, Chinese Academy of Sciences	Academic Research Group	MntModel	1.0	0,380%	0,505%	0,819%	
18/07/2012	FMISO-STD-1.0	id3 Technologies	Company	Fingerprint Matcher ISO	2.0.1	0,392%	0,592%	0,801%	
20/02/2014	FMISO-STD-1.0	Biometric System Laboratory	Academic Research Group	MCC (Baseline)	1.4	0,411%	0,602%	0,999%	
26/08/2012	FMISO-STD-1.0	Beijing Hisign Bio-info Institute	Company	FMTISO	0.24	0,439%	0,563%	0,783%	
27/01/2015	FMISO-STD-1.0	Beijing Hisign Bio-info Institute	Company	HXKJ	1.9	0,479%	0,599%	0,765%	
06/07/2014	FMISO-STD-1.0	Miguel Angel Medina-Perez	Independent Developer	M3gl	2.3.2	0,508%	0,750%	1,129%	
14/02/2014	FMISO-STD-1.0	Zengbo Xu	Independent Developer	FMISO	1.0.3	0,519%	0,700%	1,180%	
09/05/2013	FMISO-STD-1.0	Vilnius University	Academic Research Group	ABVD	1.0	0,537%	0,718%	0,985%	
26/09/2009	FMISO-STD-1.0	APRO TECHNOLOGY (BANGKOK) CO., LTD.	Company	APF_FMISO	1.1	0,582%	0,801%	1,057%	
15/07/2016	FMISO-STD-1.0	Quantobit S.L.	Company	BioCore	4.0.13	0,598%	0,768%	1,172%	
29/07/2016	FMISO-STD-1.0	Med Ghaddab	Independent Developer	test2	1.0	0,648%	1,046%	1,641%	
05/07/2016	FMISO-STD-1.0	Vsoft	Company	BioPass Finger Matching	1.8.0	0,671%	0,960%	1,331%	
09/04/2015	FMISO-STD-1.0	Arak University of Technology	Academic Research Group	Rayabin-MOC	1.54	0,751%	1,432%	2,457%	
27/10/2014	FMISO-STD-1.0	Yongliang Zhang	Independent Developer	fprMatch	0.1.1	0,797%	1,144%	1,490%	
10/10/2013	FMISO-STD-1.0	Qianbang Yang	Independent Developer	alguxntaix	1.1.3.0	0,841%	1,169%	1,620%	
06/06/2012	FMISO-STD-1.0	Muhammad Ali Jinnah, University	Academic Research	MAJU_VisPRS_FIS_Matcher	1.0	0,859%	1,778%	4,452%	

Figura B.6: Classificação dos algoritmos do projeto FVC 2006, classificação do oitavo ao vigésimo quarto colocado.

			Group						
26/12/2012	FMISO-STD-1.0	Muhammad Ali Jinnah, University	Academic Research Group	MAJU_VISPRS_FIS_algo	1.0	0,866%	1,418%	1,955%	
30/11/2010	FMISO-STD-1.0	Communik8 Ltd	Company	Authentik8	1.0	1,017%	2,475%	10,473%	
09/06/2014	FMISO-STD-1.0	Biomorf Lone	Company	LSM	1.0.1.2	1,100%	1,605%	3,463%	
01/04/2012	FMISO-STD-1.0	Robert Važan	Independent Developer	SourceAFIS	1.6	1,172%	1,797%	2,457%	
18/06/2015	FMISO-STD-1.0	Rao Tirupathi	Independent Developer	ISOSDK	0.1	1,322%	2,615%	4,242%	
08/01/2014	FMISO-STD-1.0	Avi Gazneli	Independent Developer	SpNN	1	33,050%	78,171%	84,159%	

Figura B.7: Classificação dos algoritmos do projeto FVC 2006, classificação do vigésimo quinto ao trigésimo colocado.