

FINGERPRINT DATABASE RECOGNITION THE GENETIC ALGORITHM METHOD

Afonso Vieira Davi^a; Mylene Lemos Rodrigues^b; Dogmar Antonio de Souza Junior^c

^a*Graduado em Engenharia de Computação, Universidade de Uberaba, Uberlândia,*
afonso_davi@ig.com.br

^b*Professora, Curso de Engenharia de Computação, Universidade de Uberaba, Uberlândia,*
mylene.rodrigues@uniube.br

^c*Professor, Faculdade de Engenharia Civil, Universidade Federal de Uberlândia, Uberlândia,*
souzadogmar@feciv.ufu.br

Palavras-chave: Biometria, Processamento de Imagens, Otimização, Algoritmos Genéticos.

Resumo. A tecnologia da informação está cada vez mais difundida na realidade atual, pois qualquer tipo de acesso a dados normalmente se dá por meio digital. Como consequência surge a necessidade de se criar ferramentas computacionais para proteger e garantir a segurança dessas informações, de maneira que o acesso a dados sigilosos seja permitido apenas a indivíduos autorizados. Diversas são as características biométricas utilizadas para identificação de indivíduos, dentre elas, encontram-se as impressões digitais. A escolha de uma ou outra característica está associada a diversos fatores, principalmente a sua unicidade. No caso das impressões digitais não existem duas pessoas que possuem a mesma forma com que as linhas se dispõem em seus dedos, o que atende o requisito de unicidade. Diante deste contexto foi desenvolvido o programa computacional PRID que emprega o método dos algoritmos genéticos para identificar a impressão digital de entrada no banco de dados, para tanto a função objetivo é formada a partir da distância euclidiana e o ângulo de cada minúcia em relação ao centro de massa da imagem. Os resultados obtidos nas avaliações realizadas permitiram avaliar a eficiência, desempenho e limitações do método dos Algoritmos Genéticos no processo de recuperação de impressões digitais. Pode-se concluir que o uso dos Algoritmos Genéticos para fins de reconhecimento de impressões digitais foi eficiente e eficaz para os testes realizados.

1 INTRODUÇÃO

Com o advento da internet e a popularização dos computadores na sociedade atual, faz-se necessária a implementação de programas computacionais com o intuito de proteger dados armazenados. Neste contexto a segurança da informação é uma área da computação em plena expansão e juntamente com a mesma, encontra-se a recuperação da informação, que nada mais é do que a busca eficiente de resultados em grandes bases de dados.

Existem diversas maneiras de se proteger dados, seja por criptografia, seja pelo uso de senhas pessoais. Porém atualmente há uma grande movimentação pela utilização de modelos biométricos, ou seja, sistemas que utilizem características próprias e únicas de seres humanos, para liberar ou não acesso a dados.

Dentre as características biométricas que podem ser usadas para identificar um indivíduo, destacam-se as comportamentais (por exemplo: gestos, voz, modo de andar e assinatura) e fisiológicas (impressão digital, face, geometria da mão, íris do olho, voz, entre outras). Teoricamente qualquer característica fisiológica ou comportamental humana pode ser usada para uma identificação pessoal, desde que respeite os seguintes requisitos (JAIN et al., 1997):

- universalidade: significa que todas as pessoas devem ter a característica;
- unicidade: indica que duas pessoas não podem ser a mesma em termos da característica;
- imutabilidade: o que significa que a característica deve ser invariante com o tempo;
- coletável: consiste tanto no conhecimento da característica como da tecnologia necessária para a extração das características;
- performance: refere-se tanto à precisão quanto aos recursos necessários para conseguir uma precisão aceitável na comparação dos dados biométricos (velocidade);
- aceitabilidade: refere-se à quantidade de pessoas que sentem à vontade em fornecer tal dado biométrico;
- evasão: consiste no grau de facilidade em enganar o sistema por meio de técnicas fraudulentas.

Todo sistema de recuperação de informações baseia-se no princípio de capturação de dados para armazenamento, comparação do dado consulta com os demais dados armazenados e uma resposta se o dado encontra-se ou não armazenado.

Este trabalho visa apresentar um sistema de busca biométrica através de impressões digitais utilizando o método dos Algoritmos Genéticos (AGs) para efetuar a comparação entre os dados e autorizar ou não o acesso as informações.

2 FUNDAMENTAÇÃO TEÓRICA

A impressão digital de um indivíduo consiste na representação gráfica das linhas e sulcos presentes na pele das áreas de contato, como palma das mãos, planta dos pés e ponta dos dedos (MARQUES, 2004). A Figura 1 ilustra a impressão digital do polegar de um indivíduo. As impressões digitais têm sido estudadas e utilizadas desde o início do século XX em criminalística para a identificação de indivíduos. Recentemente, com o advento dos computadores, as impressões digitais passaram a ser usadas em aplicações civis e comerciais de segurança, tais como controle de acesso, transações eletrônicas, uso de cartão de crédito e acesso a base de dados (REIS, 2003).



Figura 1: Exemplo de uma impressão digital (PRABHAKAR, 2001)

Devido à baixa probabilidade de duas impressões serem iguais, razão de 1 para $1,9 \times 10^{15}$, a impressão digital transmite segurança na sua utilização (MARQUES, 2004). Dentre as alternativas disponíveis para identificação é a que apresenta mais estudos entre todas as características biométricas. Além disso, possui a vantagem de serem de fácil coleta, alto poder de discriminação e após completamente formadas, o que ocorre aproximadamente no sétimo mês de gravidez, elas não se alteram no decorrer da vida do indivíduo exceto em caso de acidente, segundo (PRABHAKAR, 2001). No entanto, existe um problema que consiste na qualidade da própria impressão digital, que pode ser afetada por deformações permanentes ou temporárias, decorrentes desde simples cortes, ou a imersão dos dedos em água por um longo tempo (PRABHAKAR, 2001), até sujidade e gordura na ponta dos dedos. Contudo, este problema pode ser solucionado tendo em vista que um indivíduo possui dez dedos e que cada um destes tem uma impressão digital distinta de todos os outros (REIS, 2003).

Sabe-se que o desenvolvimento de um indivíduo não depende só de fatores genéticos, mas também de fatores ambientais onde este se desenvolveu. Por isso nem mesmo gêmeos univitelinos possuem impressões digitais iguais, pois a formação destas é determinada tanto pelos genes, quanto pelo micro ambiente onde cada dedo se encontra no útero (PRABHAKAR, 2001).

De acordo com (REIS, 2003), a impressão digital de um indivíduo é formada por riscas e vales presente na pele da área de contato como, por exemplo, a ponta dos dedos. Os pontos singulares “delta” e “ápice” (ou “core”, em inglês), são usados para efetuar a classificação das impressões digitais. Na Figura 2 são destacados os pontos singulares numa impressão digital.

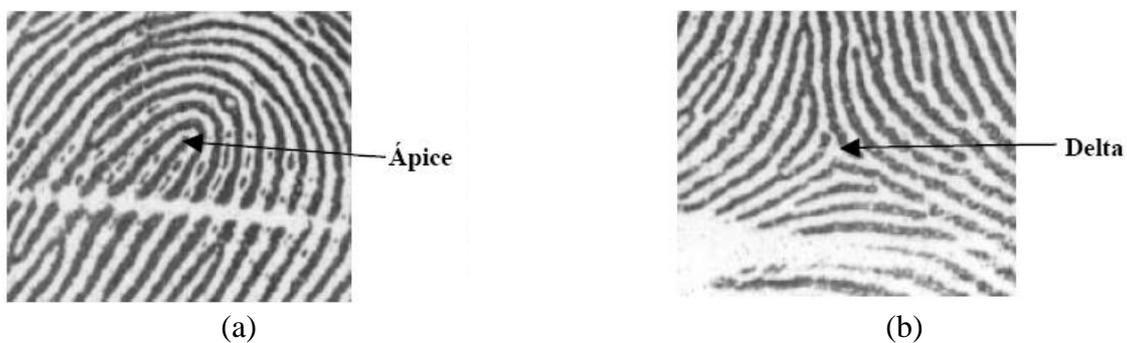


Figura 2: (a) Imagem de um ápice e (b) Imagem de um delta (MARQUES, 2004)

2.1 Método dos Algoritmos Genéticos

A própria evolução das espécies pode ser vista como um processo de otimização. Ao longo do tempo, os seres vivos tornam-se cada vez mais adaptados ao meio ambiente em constante mudança. Tal princípio foi proposto por Darwin, no século XIX, e ficou conhecido como Princípio de Seleção Natural. Logo, os AGs baseiam-se nos processos naturais de seleção, cruzamento e mutação (SOARES, 1997; SOUZA JR and GESUALDO, 2008; CASTILHO and LIMA, 2007).

Os AGs têm a vantagem de realizar buscas simultâneas em várias regiões do espaço de busca, mas apesar dessa pesquisa ser aleatória, os AGs aproveitam o conhecimento adquirido por gerações anteriores para desenvolver uma nova geração que terá uma chance maior de se aproximar da solução ótima. Com isso, os AGs trabalham com uma população e não com um único ponto, direcionando e distribuindo melhor sua busca (RIBEIRO, 2008; SOUZA JR, 2005; HOLLAND, 1975).

Os resultados que serão apresentados neste trabalho foram obtidos por meio do programa computacional PRID desenvolvido pelos autores. É importante ressaltar que o Algoritmo Genético (AG) utilizado no PRID é o mesmo desenvolvido por Souza Jr (2005) no programa computacional OTR. Para facilitar a implementação do AG, Souza Jr (2005) utilizou código real. O que não é demérito para o programa, pois o mesmo foi testado em diversas funções matemáticas e problemas físicos de otimização consagrados na literatura, e apresentou ótimos resultados.

3 METODOLOGIA

A base de dados que será utilizada neste trabalho é a base de dados do FVC2004 (Fingerprint Verification Competition) e esta disponível em <http://bias.csr.unibo.it/fvc2004/default.asp>. O FVC é a maior avaliação do mundo de tecnologia de comparação de impressão digital. É uma competição internacional para verificação de algoritmo de impressão digital, realizada a cada dois anos e organizada por laboratórios de pesquisa da Itália e Estados Unidos.

A base de dados do FVC completa é formada por quatro partes. O que diferencia cada uma das partes é o tipo de aparelho utilizado para aquisição das impressões digitais. Na parte 1 da base de dados (DB1), foi utilizado o sensor óptico “V300”, da CrossMatch. Na parte 2 (DB2), o sensor óptico “U.are.U 4000” da Digital Persona foi o escolhido para a coleta das impressões. A parte 3 (DB3) é composta por impressões adquiridas com o auxílio do sensor de varredura térmica “FingerChip FCD4B14CB” da Atmel, e, finalmente, a parte 4 da base de dados (DB4) é um conjunto de impressões digitais sintéticas criadas pelo programa “SfinGe”.

Cada parte da base de dados é formada por 100 dedos e 8 impressões digitais de cada dedo, totalizando 800 impressões digitais.

Neste trabalho foi utilizada a base de dados DB1, composta por imagens com tamanho de 640x480 pixels e resolução de 500dpi. Essa base de dados é dividida em duas partes: DB1_A formada por 800 imagens, essa será a base de dados pré-cadastrada do sistema, e DB1_B composta por apenas 80 imagens que não serão consideradas pertencentes à base de dados. Essas imagens serão utilizadas para testes de negação, ou seja, o programa não deve reconhecer tais imagens.

A execução do trabalho foi basicamente dividida em duas partes: processamento da imagem (blocos azuis) e identificação da impressão digital utilizando AGs (bloco laranja). A Figura 3 e a Figura 4 apresentam o fluxograma de funcionamento do programa computacional PRID.

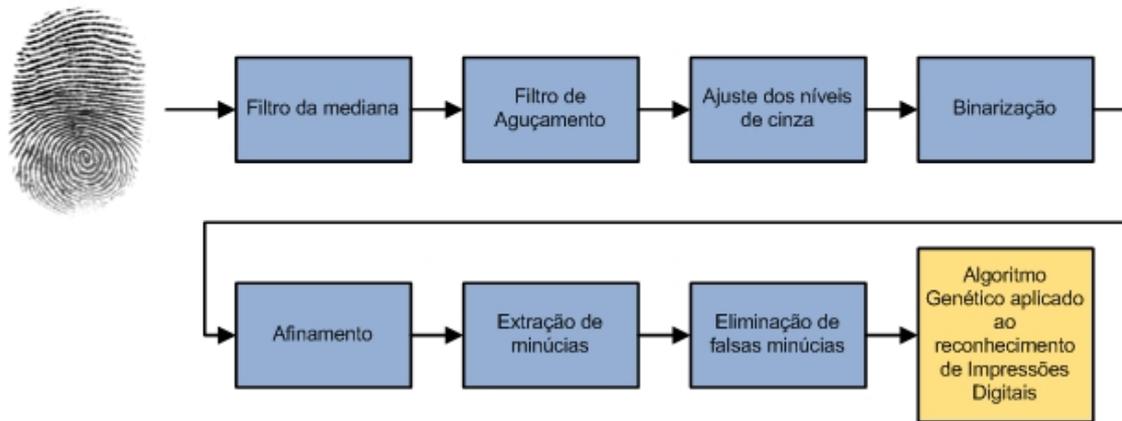


Figura 3: Fluxograma do funcionamento do sistema de autenticação

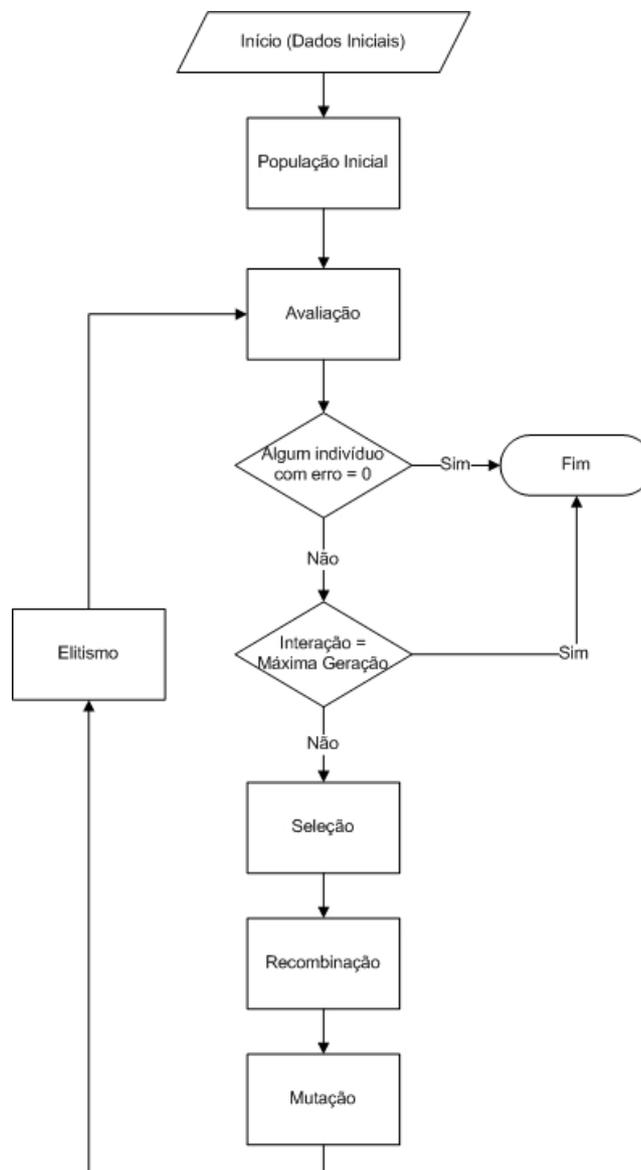


Figura 4: Fluxograma do algoritmo genético utilizado.

A etapa de processamento da imagem foi toda implementada utilizando o software Matlab, através da toolbox de processamento de imagem, enquanto a etapa de identificação das impressões digitais utilizando AGs foi desenvolvida no aplicativo Borland Delphi versão 7.0.

3.1 Filtro de suavização de ruídos

Após a aquisição da imagem é comum aparecerem pequenas degradações ou deformidades decorrentes do próprio processo de aquisição ou de algum fator externo, como poeira por exemplo. Estas imperfeições são chamadas de ruído e caso este não seja eliminado, ou pelo menos suavizado, ele pode atrapalhar o processo de extração de minúcias, podendo gerar minúcias falsas ou degradar minúcias verdadeiras, dificultando a sua detecção.

Segundo estudo realizado por (FARIA, 2005) o filtro da mediana obtém resultados muito eficientes na suavização de ruídos sem afetar detalhes da imagem e também por não possuir um custo computacional elevado e por esta razão ele foi o escolhido para realização desta tarefa. A Figura 5 apresenta um exemplo de aplicação deste filtro, em (a) imagem original; (b) imagem original com ruído do tipo sal-pimenta¹; (c) imagem resultante da aplicação do filtro.

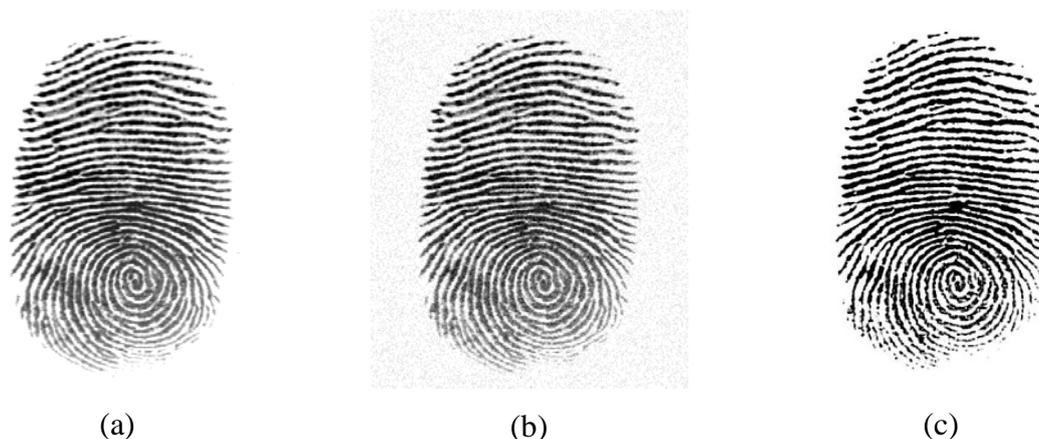


Figura 5: Exemplo de aplicação do filtro mediana

3.2 Filtro de Açuçamento

O objetivo principal do aguçamento é enfatizar detalhes finos da imagem ou realçar detalhes que tenham sido borrados (GONZALEZ; WOODS, 2003). De acordo com experimentos realizados por (FARIA, 2005) foi observado que, aplicando-se a máscara 5x5 apresentada na Figura 6 ocorre um aumento significativo do contraste e da nitidez da imagem facilitando assim o processo de binarização.

¹ Caracteriza-se pelo aparecimento de pixels de intensidade máxima (brancos) em uma parte de intensidade baixa (preta) da imagem e vice-versa, dando assim a impressão de que foi derramado sal e pimenta sobre a mesma

$$\frac{1}{3} \times$$

-1	-1	-1	-1	-1
-1	2	2	2	-1
-1	2	3	2	-1
-1	2	2	2	-1
-1	-1	-1	-1	-1

Figura 6: Máscara 5x5 usada para filtragem espacial de aguçamento

A Figura 7 mostra uma impressão digital após aplicação do filtro de aguçamento.



Figura 7: (a) imagem filtrada (mediana 3x3); (b) imagem após filtro de aguçamento

3.3 Ajuste dos níveis de cinza e binarização

O ajuste dos níveis de cinza tem por objetivo equalizar o contraste por toda a imagem a fim de evitar que falsas minúcias decorrentes da deficiência de contraste sejam geradas. Após esta etapa, foi utilizado o processo de binarização.

O processo de binarização, ou limiarização, consiste em transformar uma imagem em tons de cinza para uma imagem em preto e branco. Segundo (FARIA, 2005) dentre os diversos métodos existentes para realizar esta tarefa o que apresentou os melhores resultados foi a limiarização global de Otsu.

O método de limiarização global de Otsu consiste em escolher um limiar (nível de threshold) para minimizar a variação entre os pixels preto e branco. Assim, pixels com valor de nível de cinza menor que o nível de threshold são transformados em zero, e no caso contrário são convertidos para um. Esta operação é definida pela Eq. (1).

$$G(x, y) = \begin{cases} 0 & \text{se } f(x, y) \leq T \\ 1 & \text{se } f(x, y) > T \end{cases} \quad (1)$$

onde $G(x,y)$ representa a imagem binária resultante, $f(x,y)$ a imagem original e T o nível de threshold.

A Figura 8 apresenta o resultado da limiarização global de Otsu.



Figura 8: (a) Imagem original; (b) resultado da limiarização global de Otsu

Por fim, são removidos da imagem objetos menores que 50 pixels. Isto faz com que pequenos ‘buracos’ nas cristas da impressão digital sejam eliminados. A Figura 9 apresenta uma imagem binarizada antes e após a eliminação.



Figura 9: (a) Imagem binarizada antes da remoção dos objetos; (b) resultado

3.4 Afinamento

O afinamento é uma etapa necessária para a extração de minúcias em impressões digitais, ele pode ser dado como completo quando o número de mudanças na imagem converge para zero, ou seja, quando não há mais mudanças. Neste momento, as linhas datilares da imagem devem ter a espessura de um pixel (CASADO, 2008). A Figura 10 apresenta o resultado da operação.



(a) (b)

Figura 10: (a) Imagem binarizada; (b) Imagem afinada

No entanto, ainda é preciso remover alguns pixels errôneos da imagem. Para este trabalho aplicam-se alguns filtros morfológicos na seguinte sequência: clean, hbreak e spur (CASADO, 2008).

A Figura 11 mostra em (a) uma imagem afinada antes da aplicação dos filtros morfológicos, já em (b) é exibida a mesma imagem após a aplicação dos filtros.



Figura 11: (a) Imagem afinada; (b) Imagem após remoção de ruídos com filtros morfológicos

3.5 Extração de minúcias

Para a extração de minúcias foi utilizado o algoritmo Crossing Number (THAI, 2003). Essa escolha se deve ao fato de sua baixa complexidade computacional além de um tempo de processamento também baixo.

A Figura 12 apresenta o resultado do processo de extração de minúcias. Existe uma quantidade muito grande de minúcias, e muitas delas são consideradas falsas, ou seja, são decorrentes de algum tipo de ruído ou borrão não eliminado pela fase de pré-processamento ou até mesmo originários do processo de afinamento. É necessário, portanto, realizar um pós-processamento para a eliminação dessas minúcias falsas.

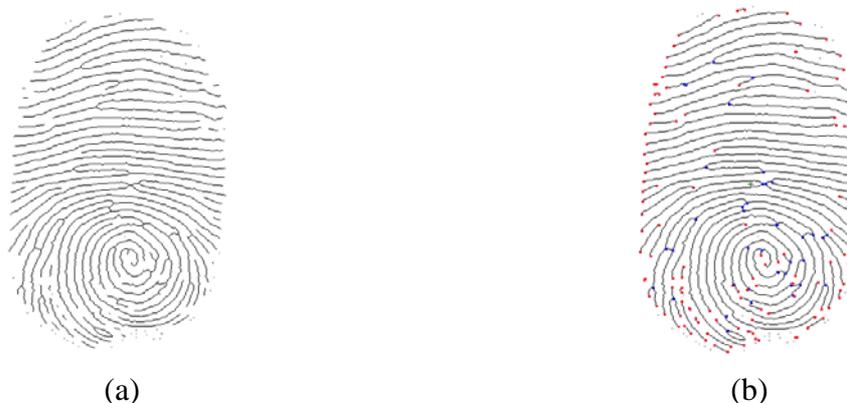


Figura 12: (a) Imagem afinada; (b) Imagem resultante do algoritmo CN onde: minúcias bifurcadas (azul), Finais (vermelho) e centro de massa (verde)

Dada a nuvem de pontos gerada pela imagem afinada (Figura 12-a) no plano cartesiano

(x,y) pode-se calcular o centro de massa desta distribuição através da Eq. (2).

$$r_{cm} = \frac{\sum m_i \cdot r_i}{\sum m_i} \quad (2)$$

3.6 Eliminação de falsas minúcias

Em um sistema de autenticação baseada no reconhecimento de impressões digitais faz-se necessário utilizar uma etapa de pós-processamento das imagens, a fim de eliminar falsas minúcias causadas por degradações na imagem. O método utilizado não remove pixels no esqueleto da imagem, mas elimina as falsas minúcias marcadas pelo algoritmo Crossing Number.

Segundo Faria (2005) apud Tico e Kuosmanen (2000) “Foi utilizado um algoritmo de pós-processamento proposto por Tico e Kuosmanen (2000). O algoritmo analisa a vizinhança de cada candidato à minúcia para decidir qual é falsa ou não”.

Para cada candidato à minúcia escolhido pela técnica do Crossing Number são feitos testes para a validação ou não desta minúcia. Os passos do algoritmo são:

1. Criar e iniciar com 0 uma máscara M de tamanho 23x23 sendo o pixel central o candidato à minúcia e cada um dos demais pixels da máscara correspondem a um pixel da imagem afinada na vizinhança 23x23;
2. Rotular com -1 o pixel central de M. Este é o pixel que corresponde ao candidato à minúcia na imagem afinada;
3. Se a minúcia é candidata a uma terminação (CN = 1):
 - a) Rotular com 1 todos os pixels na máscara M que correspondem aos pixels conectados ao candidato terminação;
 - b) Contar o número de transições de 0 para 1 (T_{01}) encontradas ao se fazer uma volta completa no sentido horário nas bordas da máscara M;
 - c) Se $T_{01}=1$ então a candidata é considerada uma terminação verdadeira.
4. Se a minúcia é candidata a uma bifurcação (CN = 3):
 - a) Fazer uma volta completa na vizinhança de 8 do pixel candidato rotulando com 1, 2, 3 respectivamente, os três componentes conexos (no formato Y) encontrados nesta volta;
 - b) Para cada pixel rotulado de 1 deve-se procurar seus pixels conexos dentro da máscara M, isto é, os pontos pretos que formam as linhas datílares da impressão digital. Fazer a rotulação até o final de M ou encontrar outro valor de rótulo. Repetir este passo para o rótulo 2 rotulando os componentes conexos com 2 e também para o rótulo 3, rotulando seus conexos com 3;
 - c) Contar o número de transições de 0 para 1 (T_{01}), 0 para 2 (T_{02}) e de 0 para 3 (T_{03}) encontradas quando feita uma volta completa no sentido horário nas bordas de M;
 - d) Caso $T_{01}=1, T_{02}=1$ e $T_{03}=1$, então validar a minúcia candidata como uma bifurcação verdadeira.

Alguns exemplos de falsas minúcias eliminadas por esse algoritmo são mostradas na Figura 13.

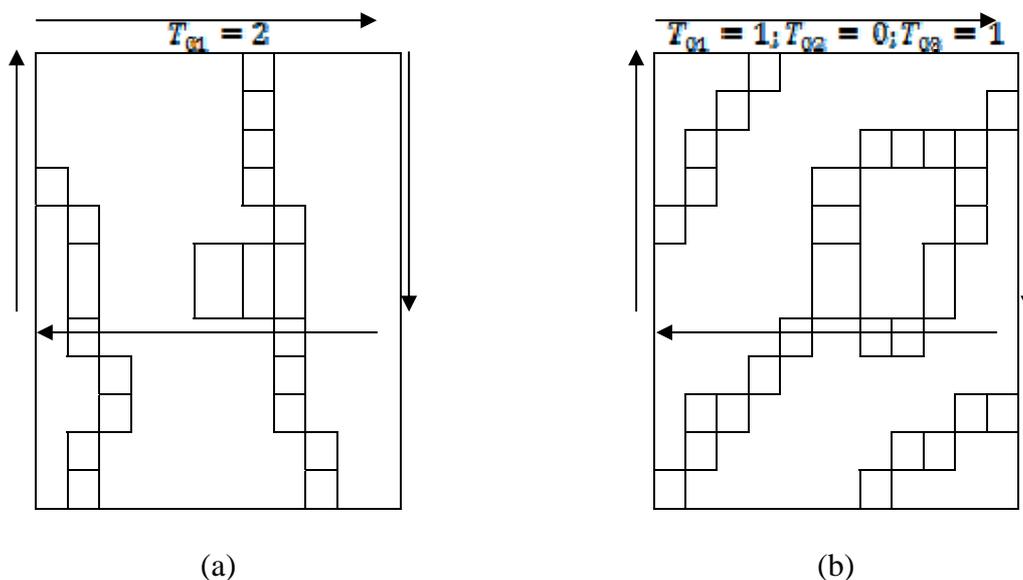


Figura 13: Exemplos de falsas minúcias canceladas. (a) espigada; (b) buraco

A Figura 14 apresenta um exemplo de impressão digital pós-processada. A quantidade de minúcias que permaneceram na imagem é menor que 20%. Note ainda que as minúcias localizadas nas extremidades da imagem também são eliminadas pelo pós-processamento.

Para cada impressão digital é gerado um vetor de minúcias, contendo nas duas primeiras posições as coordenadas cartesianas do centro de massa e nas próximas posições as coordenadas cartesianas e tipo de cada minúcia validada pelo pós-processamento. Cada um desses vetores é salvo, compondo assim o template. As minúcias são salvas com a coordenada x sempre crescente, ou seja, as minúcias são salvas de cima para baixo e da esquerda para a direita.



Figura 14: (a) Imagem resultante do Crossing Number (176 minúcias); (b) imagem pós-processada (32 minúcias)

3.7 Operadores genéticos

A seguir são apresentados os operadores genéticos empregados no AG do programa computacional PRID.

3.7.1 Seleção pelo método da roleta

O método da roleta é o mais simples e consiste em associar a cada indivíduo um número

que representa a probabilidade deste indivíduo a passar para a próxima geração em função do seu grau de aptidão como solução para o problema em estudo. Nesta etapa, os indivíduos melhor adaptados ao problema receberão um valor de probabilidade maior, isto representa dizer que eles terão maior chance de serem selecionados para a próxima geração, enquanto que os indivíduos com menor probabilidade tendem a desaparecer da população. Em seguida determina-se a probabilidade acumulada de cada indivíduo, de maneira que, o último indivíduo terá probabilidade acumulada igual a 1. O próximo passo é gerar aleatoriamente “Pop” números aleatoriamente entre 0 e 1. Cada número gerado apontará para um indivíduo da roleta conforme ilustrado na Figura 15.

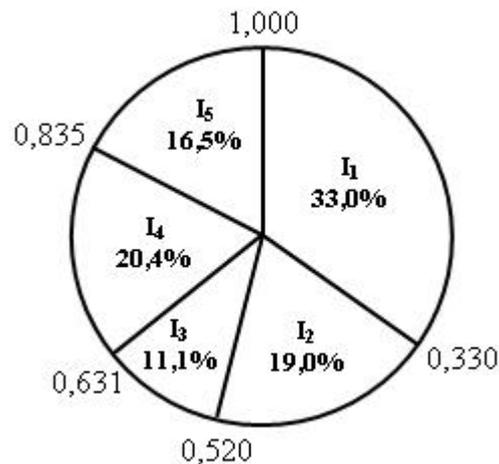


Figura 15: Exemplo de seleção pelo método da roleta

3.7.2 Recombinação

Nesta etapa combina-se partes de cromossomos de dois indivíduos diferentes (pais), gerando assim um novo indivíduo com características de ambos os pais (filho). Esse procedimento é usado em geral, para garantir a diversidade da população (SOUZA JR, 2005). O cruzamento é controlado pela taxa de cruzamento P_c . Uma taxa de cruzamento muito alta aumenta o espaço de busca porém dificulta a convergência para um ótimo local. Já uma taxa muito baixa pode ocasionar uma perda de tempo computacional por abranger regiões não promissoras no espaço de busca (SOUZA JR, 2005). Dentre os diversos modelos de operadores para recombinação disponíveis na literatura, foi escolhido implementado o método da recombinação discreta.

3.7.3 Mutação

O operador mutação tem por objetivo garantir que diversas alternativas sejam exploradas, mantendo assim uma boa abrangência do espaço de busca. Ele é controlado por um padrão fixo denominado P_m . Esta probabilidade refere-se ao total de bits da população que deverá ser mutacionada (RIBEIRO, 2008).

Além de permitir que o algoritmo faça buscas em diferentes regiões do espaço o processo de mutação protege o algoritmo da perda de material genético potencialmente útil pela aplicação dos operadores reprodução e recombinação (SOUZA JR, 2005).

3.7.4 Elitismo

Quando se passa de uma geração para outra pode acontecer do indivíduo melhor adaptado não ser selecionado. Para evitar este problema e garantir que o indivíduo melhor adaptado ao problema em cada geração passe para a próxima geração implementou-se o método do elitismo.

3.8 Função de objetivo

A Eq. (3) apresenta a função objetivo implementada no AG. Ressalta-se que a função aptidão é a função inversa da função objetivo.

$$F(i) = \sqrt{\sum_{j=1}^N \frac{\sum_{z=1}^T \text{Erro}_z}{(N_{\text{idéntica}}^2 + N_{\text{verdadeira}})}} \quad (3)$$

$$\text{Erro}_z = |(d_{\text{entrada}} - d_{\text{template}}) \cdot (\theta_{\text{entrada}} - \theta_{\text{template}})|$$

$$d_{\text{entrada}} = \sqrt{(CM_X - A_X)^2 + (CM_Y - A_Y)^2}$$

Onde:

CMx = coordenada X do centro de massa da imagem;

CMy = coordenada Y do centro de massa da imagem;

Ax = coordenada X da minúcia;

Ay = coordenada Y da minúcia;

θ_{entrada} = ângulo da minúcia analisada em relação ao centro de massa da imagem analisada, calculado conforme Figura 15;

θ_{template} = ângulo da minúcia comparada (do banco de dados) em relação ao centro de massa da imagem armazenada no banco de dados (template), calculado também conforme Figura 16;

d_{entrada} = distância euclidiana da minúcia analisada em relação ao centro de massa da imagem analisada;

d_{template} = distância euclidiana da minúcia comparada em relação ao centro de massa da imagem armazenada no banco de dados (template);

N = Número de minúcias analisadas;

T = Número de minúcias encontradas na vizinhança de 80x80 da imagem do banco de dados (template);

N_{idéntica} = Número de minúcias idênticas encontradas no processo de extração (Erro = 0);

N_{verdadeira} = Número de minúcias com Erro > 0;

i = indivíduo.

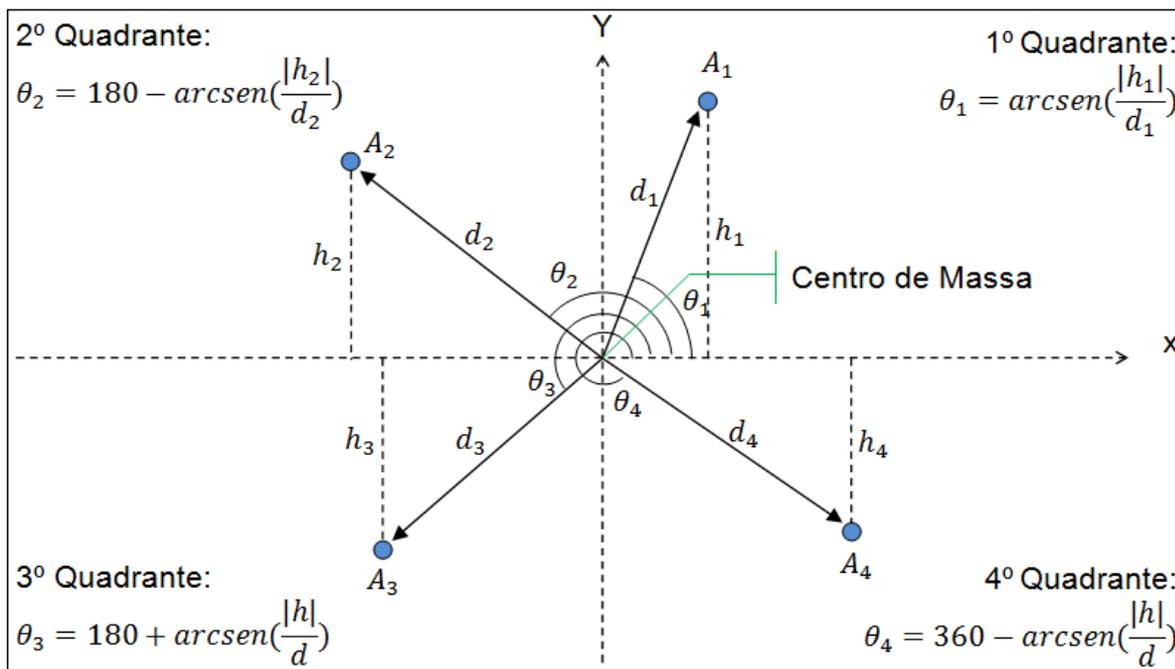


Figura 16: Cálculo do ângulo da minúcia em relação ao centro de massa

4 RESULTADOS

Para a análise de desempenho do AG no reconhecimento de impressões digitais foi selecionado aleatoriamente as vinte imagens apresentadas na Tabela 1. Lembrando que as imagens pertencentes à base de dados DB1_A estão pré-cadastradas, ou seja, elas devem ser aceitas pelo sistema, já as imagens de DB1_B não foram pré-cadastradas, logo elas devem ser negadas pelo sistema, pois, são “intrusas”.

	ID	Base de dados		ID	Base de dados
1	7_4	DB1_A	11	104_4	DB1_B
2	43_5	DB1_A	12	109_5	DB1_B
3	75_3	DB1_A	13	110_4	DB1_B
4	17_3	DB1_A	14	37_5	DB1_A
5	46_7	DB1_A	15	66_5	DB1_A
6	100_4	DB1_A	16	1_3	DB1_A
7	20_2	DB1_A	17	30_6	DB1_A
8	74_5	DB1_A	18	107_6	DB1_B
9	1_6	DB1_A	19	104_1	DB1_B
10	102_6	DB1_B	20	27_1	DB1_A

Tabela 1: Imagens selecionadas aleatoriamente para realização dos testes.

onde: QTM: Quantidade total de minúcias; ID: Impressão Digital; TC: Tempo de Comparação; QMA: Quantidade de minúcias analisadas; IE: Imagem encontrada (imagem mais semelhante segundo o algoritmo); FA: Falsa Aceitação (IDs que não pertencem á base de dados mas o sistemas as reconhece como se fossem); FR: Falsa Rejeição (IDs que pertencem à base de dados, mas o sistema não as reconhece).

As métricas VFO (Valor da Função Objetivo) e PMI (Percentual de minúcias Idênticas em relação a QMA) são utilizadas no algoritmo de comparação. Elas foram obtidas empiricamente, onde após alguns testes com imagens semelhantes, transladadas, rotacionadas e as que foram adicionados ruídos. Foi constatado que se o resultado do VFO estiver entre 0 e 1 e o PMI maior que 50% a impressão digital foi encontrada no banco de dados. Quanto mais próximo de zero estiver o VFO e mais próximo de 100% estiver o PMI mais semelhante é a impressão digital que está sendo identificada com a imagem que está armazenada na base de dados.

Para todos os testes realizados a quantidade de minúcias a ser analisada pelo algoritmo é de 50% do total de minúcias da imagem de entrada. Os testes para análise do desempenho do AG foram divididos em três sequências. A Tabela 2 apresenta os dados de entrada do AG para as três sequências.

	Sequência 1	Sequência 2	Sequência 3
População	100	20	50
Gerações	1000	40	100
Taxa de mutação	0,3	0,3	0,3
Taxa de recombinação	0,9	0,9	0,9

Tabela 2: Dados de entrada do AG.

Os resultados obtidos para as sequências de testes 1, 2 e 3 estão apresentados na Tabela 3, Tabela 4 e Tabela 5, respectivamente.

	ID	IE	TC (s)	VFO	QTM	QMA	PMI	FA	FR
1	7_4	7_4	0,180	0	32	16	100%		
2	43_5	43_5	0,138	0	37	18	100%		
3	75_3	75_3	0,200	0	22	11	100%		
4	17_3	17_3	0,179	0	13	6	100%		
5	46_7	46_7	0,071	0	42	21	100%		
6	100_4	100_4	0,061	0	33	16	100%		
7	20_2	20_2	0,103	0	31	15	100%		
8	74_5	74_5	0,169	0	54	27	100%		
9	1_6	1_6	0,189	0	45	22	100%		
10	102_6	37_5	6,719	0,29	46	23	100%	X	
11	104_4	88_5	0,981	0	32	16	100%	X	
12	109_5	88_4	4,557	0,348	38	19	100%	X	
13	110_4	84_3	4,751	0,283	46	23	100%	X	
14	37_5	37_5	0,124	0	68	34	100%		
15	66_5	66_5	0,126	0	29	14	100%		
16	1_3	1_3	0,103	0	52	26	100%		
17	30_6	30_6	0,268	0	96	48	100%		
18	107_6	63,5	7,252	4,644	89	44	61%		
19	104_1	13_3	3,439	0,595	26	13	100%		X
20	27_1	27_1	0,047	0	34	17	100%		

Tabela 3: Resultados apresentados pela sequência de teste 1.

	ID	IE	TC (s)	VFO	QTM	QMA	PMI	FA	FR
1	7_4	9_3	0,108	11,27	32	16	68%		X
2	43_5	43_5	0,045	0	37	18	100%		
3	75_3	30_5	0,099	62,18	22	11	45%		X
4	17_3	58_3	0,077	620,95	13	6	66%		X
5	46_7	46_7	0,037	0	42	21	100%		
6	100_4	42_3	0,103	128,57	33	16	31%		X
7	20_2	20_2	0,062	0	31	15	100%		
8	74_5	5_3	0,106	13,23	54	27	66%		X
9	1_6	53_6	0,115	6,92	45	22	77%		X
10	102_6	96_2	0,117	2,03	46	23	73%		
11	104_4	68_2	0,096	12,68	32	16	62%		
12	109_5	50_4	0,120	35,38	38	19	52%		
13	110_4	83_3	0,119	37,05	46	23	43%		
14	37_5	37_5	0,070	0	68	34	100%		
15	66_5	66_5	0,072	0	29	14	100%		
16	1_3	15_6	0,118	157,02	52	26	42%		X
17	30_6	30_6	0,156	0,02	96	48	100%		
18	107_6	64_6	0,142	121,52	89	44	34%		X
19	104_1	59_1	0,072	112,25	26	13	69%		X
20	27_1	27_1	0,066	0	34	17	100%		

Tabela 4: Resultados apresentados pela sequência de teste 2.

	ID	IE	TC (s)	VFO	QTM	QMA	PMI	FA	FR
1	7_4	7_4	0,063	0	32	16	100%		
2	43_5	43_5	0,076	0	37	18	100%		
3	75_3	75_3	0,071	0	22	11	100%		
4	17_3	17_3	0,120	0	13	6	100%		
5	46_7	46_7	0,062	0	42	21	100%		
6	100_4	100_4	0,074	0	33	16	100%		
7	20_2	20_2	0,142	0	31	15	100%		
8	74_5	74_5	0,290	0	54	27	100%		
9	1_6	1_6	0,132	0	45	22	100%		
10	102_6	50_5	0,385	1,14	46	23	100%		
11	104_4	76_3	0,239	13,38	32	16	81%		
12	109_5	93_6	0,333	1,73	38	19	94%		
13	110_4	96_4	0,258	16,91	46	23	69%		
14	37_5	37_5	0,226	0	68	34	100%		
15	66_5	66_5	0,050	0	29	14	100%		
16	1_3	1_3	0,065	0	52	26	100%		
17	30_6	30_6	0,379	0	96	48	100%		
18	107_6	64_5	0,439	48,92	89	44	36%		
19	104_1	28_1	0,280	1,17	26	13	100%		
20	27_1	27_1	0,110	0	34	17	100%		

Tabela 5: Resultados apresentados pela sequência de teste 3.

As imagens destacadas em azul são aquelas pertencentes à base de dados e que foram corretamente reconhecidas pelo sistema, as vermelhas são também pertencentes à base de dados, no entanto, o sistema não as identificou ou as identificou como outra imagem. Já as verdes são imagens não pertencentes à base de dados e que foram corretamente não reconhecidas pelo sistema, e por fim, as imagens laranja são aquelas também não pertencentes à base, porém foram erroneamente reconhecidas. A Tabela 6 apresenta um comparativo entre as três sequências de testes.

	Média do tempo de consulta (TC)	TFA	TFR	Eficácia
Sequência 1	1,483 s	83,33%	0%	75%
Sequência 2	0,095 s	0%	37,5%	55%
Sequência 3	0,190 s	0%	0%	100%

Tabela 6: Comparativo entre as sequências de testes.

Pela Tabela 6 pode-se notar como os parâmetros da Sequência 1 tornam o sistema mais lento e com uma aceitação muito grande de imagens falsas. Já a Sequência 2 é a mais rápida de todas, no entanto, possui baixa eficácia, ou seja, não reconhece imagens cadastradas e não nega imagens falsas. Enquanto a Sequência 3 possui uma eficiência (velocidade de processamento) aproximadamente duas vezes maior que a Sequência 2 e quase 8 vezes melhor que a Sequência 1. Além disso, teve excelente eficiência, aceitando todas as imagens pré-cadastradas e rejeitando todas as não pertencentes à base de dados analisando apenas 50% das minúcias encontradas pelo pós-processamento.

Através destes resultados, chegou-se à conclusão de que os melhores parâmetros para a execução do AG no reconhecimento de impressões digitais são os da Sequência 3. Estes resultados demonstram um desempenho satisfatório do sistema na Sequência 3, pois nenhuma das verificações levou mais de meio segundo para ser executada, ficando na excelente média de 190 milésimos de segundo.

Comparando com as dissertações de Faria (2005) e Silvia (2001), este sistema obteve um melhor tempo de processamento. No primeiro, o tempo resultante da comparação de uma Impressão Digital com toda a base de dados (apenas 81 imagens cadastradas) leva em média de 10 à 12 segundos, utilizando uma máquina com Pentium 4 de 2,5 GHz com 256 Mb de memória RAM. Já em (SILVIA, 2001) o processo de verificação de uma impressão digital com sua base de dados composta por 196 imagens gasta em média 19,4 segundos utilizando um Pentium 2 de 300 Ghz e 64 Mb de RAM.

È interessante observar que quando ocorre uma FA ou FR o tempo de processamento é maior, principalmente no caso das FAs. Isso ocorre porque a disposição das minúcias na impressão analisada é bastante similar à de alguma impressão da base de dados e, como o AG trabalha para minimizar o erro buscando sempre as minúcias mais similares, conseqüentemente, a cada interação o erro resultante é menor que o anterior. Isso faz com que o AG execute até a última geração, o que não ocorre quando ele encontra uma impressão idêntica pois, quando o erro é zero ele termina a seleção independentemente da geração em que o mesmo se encontra.

5 CONCLUSÃO

Neste trabalho foi apresentado um sistema de verificação de indivíduos através de impressões digitais baseado em otimização genética. Para a construção deste sistema de verificação foi necessário realizar uma vasta pesquisa sobre técnicas de processamento de

imagens onde, depois de diversas tentativas e testes, foi decidido empregar as técnicas apresentadas por Faria (2005) com algumas adaptações para a posterior utilização do AG no processo de comparação das impressões digitais. Tal escolha deve-se ao fato de que seus experimentos apresentavam resultados satisfatórios e uma implementação relativamente simples.

Através dos resultados apresentados chegou-se à conclusão de que os melhores parâmetros para a execução do algoritmo genético no reconhecimento de impressões digitais são os da Sequência 3, ou seja, uma população inicial de 50 indivíduos em 100 gerações, taxa de mutação igual a 0,3% e taxa de mutação igual a 0,9%. Utilizando tais valores para os parâmetros citados podemos dizer que a verificação de impressões digitais utilizando o método do Algoritmo Genético apresentou uma excelente eficiência e eficácia. Em termos de eficiência, o sistema levou em média, apenas 190 milésimos de segundo para encontrar a impressão digital correta dentre as 800 imagens existentes na base de dados.

Vale ressaltar que novos testes estão sendo realizados com impressões digitais transladas, rotacionadas e com presença de ruído.

REFERÊNCIAS

- Casado, R. S. *Extração de Minúcias em imagens de Impressões Digitais*. 2008. 116 f. Dissertação (Mestrado em Engenharia Elétrica) – Escola de Engenharia de São Carlos/USP, Universidade de São Paulo, São Carlos, 2008. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/18/18152/tde-15102008-135808/>>. Acesso em: 15 set. de 2009.
- Castilho, V. C. and Lima, M. C. V. Cost optimisation of lattice-reinforced joist slabs using genetic algorithms. *Structural Concrete*, 111-118, 2007.
- Faria, D. R. *Reconhecimento de impressões digitais com baixo custo computacional para um sistema de controle de acesso*. 2005. 100 f. Dissertação (Mestrado) - Programa de Pós-graduação em Informática, Setor de Ciências Exatas, Universidade Federal do Paraná, Curitiba, 2005. Disponível em: <<http://hdl.handle.net/1884/6784>>. Acesso em: 10 fev. 2008.
- Gonzalez, R. C. and Woods, R. E. *Processamento de Imagens Digitais*. São Paulo: Edgard Blücher Ltda, 2000.
- Holland, J. H. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI, 1975.
- Jain, A.; Hong, L.; Pankanti, S. and Bolle, R. *An Identity Authentication System Using Fingerprints*. 199-. Disponível em: <<http://citeseer.ist.psu.edu/jain97identity.html>>. Acesso em: 03 dez. 2007.
- Marques, A. C. P. B. *Extração de Minúcias em Imagens de Impressões Digitais utilizando Redes Neurais*. 2004. 183 f. Dissertação (Mestrado em Ciência da Computação) - Núcleo de Computação Eletrônica, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2004.
- Prabhakar, S. *Fingerprint Classification and Matching Using a Filterbank*. 2001. 259 f. Dissertação (Doutorado) - Department of Computer Science And Engineering, Michigan State University, Michigan, 2001. Disponível em: <<http://citeseer.ist.psu.edu/482312.html>>. Acesso em: 15 fev. 2008.
- Reis, C. M. S.. *Autenticação com Impressão Digital*. 2003. 78 f. Relatório de Licenciatura, Curso de Engenharia de Sistemas de Telecomunicações e Eletrônica, Departamento de Engenharia de Eletrônica e Telecomunicações e de Computadores, Instituto Superior de Engenharia de Lisboa, Lisboa, 2003. Disponível em:

<<http://www.deetc.isel.ipl.pt/comunicacoesep/disciplinas/pfc/fingerprint/files/carlos.pdf>>.

Acesso em: 03 dez. de 2007.

Ribeiro, L. M. P. *Otimização e dimensionamento de treliças planas de madeira empregando o método dos algoritmos genéticos*. 2008. 102 f. Dissertação (Mestrado em Engenharia Civil) - Faculdade de Engenharia Civil, Universidade Federal de Uberlândia, 2008.

Soares, G. L. *Algoritmos Genéticos: Estudo, Novas Técnicas e Aplicações*. 1997. 145 f. Programa de pós-graduação em Engenharia Elétrica, Centro de Pesquisa e Desenvolvimento em Engenharia Elétrica - Universidade Federal de Minas Gerais, 1997. Disponível em: <www.cpdee.ufmg.br/~joao/TesesOrientadas/VAS1997_1.pdf>. Acesso em: fev. 2008.

Souza Jr, D. A. *Otimização pelo Método dos Algoritmos Genéticos e Dimensionamento de Estruturas Tubulares Metálicas espaciais com Barras Cruzadas para Cobertura*. 2005. 125f. Tese (Doutorado em Engenharia Mecânica) – Faculdade em Engenharia Mecânica, Universidade Federal de Uberlândia, 2005.

Souza Jr, D. A. and Gesualdo, F. A. R., *Distribuição de montantes e diagonais baseada na minimização do volume de madeira de treliças planas*. Revista Madeira: Arquitetura e Engenharia, v. 21, p. 192-204, 2007.

Thai, R. *Fingerprint Image Enhancement and Minutiae Extraction*. In: School of Computer Science and Software Engineering, The University of Western Australia, 2003. Disponível em: <www.csse.uwa.edu.au/~pk/studentprojects/raymondthai/RaymondThai.pdf>. Acesso em: 28 ago. 2009.