

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO



FEUP

Reconhecimento Biométrico Baseado na Geometria da Mão

Hélder José da Silva Matos

Mestrado Integrado em Engenharia Electrotécnica e de Computadores, Major
Automação

Orientador: Prof. Doutor Jaime dos Santos Cardoso

Co-orientador: Mestre Filipe Magalhães

Janeiro de 2011

Resumo

Os sistemas de reconhecimento baseados na geometria da mão estão entre as mais antigas ferramentas biométricas utilizadas para a autenticação automática de uma pessoa. Várias patentes deste tipo de sistemas de reconhecimento foram já emitidas e vários dispositivos de controlo de acesso têm sido fabricados e comercializados desde finais dos anos 70, sendo usados, por exemplo, em oficinas nucleares, aeroportos e hotéis ao longo dos últimos 40 anos.

Sistemas de reconhecimento baseados na geometria da mão foram usados para fins de segurança no departamento de energia da ONI (Office of Naval Intelligence). No entanto, a biometria baseada em características extraídas da mão tem alcançado um enorme interesse na área académica, principalmente devido ao progresso da investigação em visão artificial e processamento e análise de imagem. Os sistemas de reconhecimento baseados na mão apresentam elevada confiança, baixo custo e são dispositivos simples de interagir para o utilizador. De um modo geral, esta é uma solução viável para aplicações de controlo em grupos de dimensões não muito elevadas.

Outras modalidades concorrentes são, por exemplo, o reconhecimento biométrico da face, íris, impressão digital e da retina. Em contraste com outras técnicas, a biometria baseada nas características da mão oferece algumas vantagens: a aquisição de dados é económica, podendo ser realizada através de sistemas comerciais de baixa resolução como scanners ou câmaras e o seu processamento é relativamente simples. Em segundo lugar, as pessoas preferem sistemas de acesso menos intrusivos como os baseados na mão ou face, pois sistemas baseados na íris ou nas veias são dos mais invasivos e as impressões digitais constituem informação privada de cada pessoa dado ser parte do seu registo civil. Assim, as características físicas da mão são consideradas como não invasivas e mais convenientes de utilização para sistemas de reconhecimento. Em terceiro lugar, os sistemas de acesso baseados na mão são adequados para uso interior e ao ar livre podendo funcionar correctamente em condições meteorológicas extremas e em deficientes condições de iluminação. Em quarto lugar, este tipo de sistemas apresentam-se como sendo os mais económicos dada a sua aquisição ser feita com equipamentos básicos, e de muito baixo custo quando comparados com os restantes. Finalmente, a informação biométrica da mão mostra ser bastante fiável, com elevado sucesso no reconhecimento de pessoas entre as populações na ordem de várias centenas. Este tipo de reconhecimento biométrico também pode ser utilizado para reforçar a segurança do comércio electrónico e de aplicações bancárias através da sua integração nos sistemas convencionais, utilizando códigos PIN ou senhas como complemento.

A biometria baseada em características da mão fundamenta-se num conjunto de medidas que podem ser extraídas recorrendo a alguns pontos característicos da mão, as cinco pontas dos dedos e os quatro vales entre eles. Um dos grandes obstáculos destes sistemas reside na detecção destes pontos e, dada a dificuldade nessa detecção, muitas das vezes são impostas restrições na fase aquisição de modo a facilitar a mesma. Estas restrições causam desconforto ao utilizador e limitam a aplicação deste tipo de sistemas. Neste trabalho foram desenvolvidas técnicas com o intuito de ultrapassar as limitações impostas pelos sistemas actuais. Um dos primeiros passos passou pela criação de uma Base de Dados livre de restrições durante a aquisição. Posteriormente foi imple-

mentado um método de detecção dos pontos característicos da mão (pontas dos dedos e vales). Para as pontas dos dedos foi desenvolvida uma metodologia baseada no esqueleto da mão, enquanto para os vales foi feita uma análise da curvatura do contorno da mão. Um grande obstáculo encontrado nesta fase residiu na binarização da mão em relação ao *background*, que tende a falhar quando esta se encontra com os dedos juntos. Identificados os pontos, foi feita uma extração das características necessárias à fase de identificação. Nesta fase, alguns algoritmos de classificação foram implementados sendo que, os que mostraram melhores resultados apresentaram uma taxa de reconhecimento de 76% com uma taxa de intrusão de 8%.

Abstract

Hand geometry based systems are among the oldest used for automatic person identification. Various patents for these types of systems have been granted and since the 1970s various devices to control access have been manufactured and marketed. Over the last 40 years these devices have been used in head offices, airports and hotels. Systems based on the hand geometry were also used for security reasons, for example, in the department of energy of the Office of Naval Intelligence.

However, biometry based on hand features has received enormous academic interest. This is principally due to the advances in research in the areas of computer vision and the image processing and analysis. Palmprint recognition systems offer high consumer confidence, low costs and they are devices easy to use. In general, these systems present a viable solution for access control of reasonable sized groups.

Other automatic person authentication systems are, for example, biometric recognition of the face, iris, fingerprint and retina. In contrast to other techniques, hand geometry does offer some advantages: with regards to data acquisition it is very economical and can be performed with low resolution commercial systems such as scanners or cameras; and the processing of this information is relatively simple. Furthermore, people prefer access control systems that are less intrusive, such as those based on the hand or the face. Systems that are based on the iris and the veins are among the most intrusive and fingerprints are part of each person's private information for civil registration. Therefore, the physical characteristics of the hand or palm are considered less invasive and more convenient to use for identity authentication. Furthermore, systems based on the hand can be used both inside buildings and outdoors; they can function in extreme weather conditions and weak light conditions. This system is also the most economical because it uses basic systems for data acquisition that are much cheaper when compared with other systems. Finally, the biometric information of the hand has been shown to be quite reliable and has a high success rate for identity authentication among hundreds. This type of identity authentication can also be used to reinforce security in electronic commerce and banking if integrated into conventional systems using PIN codes and passwords to support it.

Hand features are based on the extraction of certain points on the hand; there are five points on the fingertips and four points in the gaps between them. One of the challenges is the detection of these points as it is difficult and very often methods have to be implemented during the acquisition stage to make this easier. These restrictions can be uncomfortable for the user and can limit the application of this type of system. In this study, techniques to overcome the difficulties and limitations of current systems have been developed. One of the first steps was to create a Database free of the restrictions of data acquisition. Prior to this, a detection method for the five points (on the fingertips and in the valleys) of the hand had been implemented. With regards to the points on the fingertips, a methodology was developed based on the hand skeleton of the hand and for the valleys points an analysis was performed on the curvature of the hand contour. One difficulty was in the segmentation of the hand in relation to the background, which often fails if the fingers are closed.

Once the points had been located, the necessary features for authentication were extracted. In this stage some classification algorithms were implemented. Those showing the best results presented a Genuine Acceptance Rate (GAR) of 76% and 8% of False Acception Rate (FAR).

Agradecimentos

Ao Professor Doutor Jaime dos Santos Cardoso, orientador da FEUP e ao Mestre Filipe Magalhães, orientador da instituição, pelos conhecimentos transmitidos, apoio e disponibilidade sempre demonstrados e pela orientação e sugestões importantes ao longo de toda a dissertação.

Ao Mestre Hélder Oliveira, pelo apoio incansável, integração no grupo e excelente ambiente de trabalho.

Ao INESC Porto pela oportunidade de realização deste trabalho em ambiente institucional e nomeadamente à Unidade de Telecomunicações e Multimédia pelo acolhimento e excelentes condições de trabalho providenciadas.

Aos elementos pertencentes à Unidade de Telecomunicações e Multimédia e Serviço de Comunicação do INESC Porto, pela colaboração na criação da Base de Dados.

A toda a minha família, pelo apoio, confiança e motivação. Principalmente, aos meu avôs, João Oliveira Costa e José Matos Pereira, por me guiarem neste caminho e por estarem sempre a meu lado. Obrigado avôs!

Aos meus pais, pelo apoio incondicional que demonstraram, por toda a confiança depositada, pelos conselhos e pelo sacrifício que fizeram para me proporcionar esta graduação.

Ao meu irmão, Fábio, pela verdadeira amizade e apoio ao longo destes anos.

À Catarina, por todo o carinho, apoio e confiança nas minhas capacidades para a conclusão deste trabalho.

Aos meus amigos, Bruno Oliveira, Nuno Machado, Hugo Vasconcelos, Luís Vigário e Raquel Pereira, pela amizade, companheirismo e apoio ao longo destes anos académicos.

Hélder Matos

*”Whatever you do in life, will be insignificant.
But it’s very important that you do it, because nobody else will”*

Mohandas Karamchand Gandhi

Conteúdo

1	Introdução	1
1.1	Enquadramento	1
1.2	Motivação	1
1.3	Objectivos	2
1.4	Descrição do Trabalho Proposto e Metodologia Utilizada	3
1.5	Contribuições	5
1.6	Estrutura do Documento	6
2	Reconhecimento Biométrico	7
2.1	Sistemas Biométricos	7
2.2	Reconhecimento Biométrico Baseado na Geometria da Mão	11
2.2.1	Modos de Aquisição	11
2.2.2	Detecção de Pontos, Extração de Características e Resultados	13
2.3	Limitações dos Sistemas Baseados na Geometria da Mão	21
3	Aquisição de Imagens para o Sistema de Reconhecimento	22
3.1	Introdução	22
3.2	Tipo de Aquisição	22
3.3	Base de Dados Criada	23
3.4	Conclusão	24
4	Pré-processamento de Imagem	26
4.1	Introdução	26
4.2	Segmentação Binária	27
4.2.1	Introdução	27
4.2.2	Métodos e Resultados	29
4.2.3	Sumário	47
4.3	Extração de Contornos	47
4.3.1	Introdução	47
4.3.2	Métodos e Resultados	47
4.3.3	Sumário	52
4.4	Detecção do ângulo de orientação e alinhamento das imagens	53
4.4.1	Introdução	53
4.4.2	Métodos e Resultados	54
4.4.3	Resultados	57
4.4.4	Sumário	59
4.5	Conclusão	59

5	Detecção de Pontos Característicos	61
5.1	Introdução	61
5.2	Métodos e Resultados	61
5.2.1	Transformação para Coordenadas Polares e Análise da Curvatura	61
5.2.2	Esqueleto da Mão	65
5.2.3	<i>Polígono Convexo da Mão</i>	72
5.3	Conclusão	74
6	Definição e Extracção das Características para o Reconhecimento	78
6.1	Introdução	78
6.2	Métodos e Resultados	79
6.2.1	Extracção da Área, Perímetro e Excentricidade	79
6.2.2	Extracção do Comprimento do Dedo	80
6.2.3	Extracção da Largura do Dedo	81
6.2.4	Extracção da Região da Ponta do Dedo	83
6.3	Estudo dos Atributos	83
6.4	Conclusão	90
7	Identificação e Reconhecimento	91
7.1	Introdução	91
7.2	Métodos e Resultados	92
7.2.1	<i>Feature Matching</i>	92
7.2.2	<i>SVM - Support Vector Machine</i>	97
7.2.3	Árvores de Decisão	99
7.3	Conclusão	101
8	Conclusão Geral	103
9	Trabalho Futuro	105
A	Hand Images Data Base	107
B	Hand Characteristic Points Detection Competition	109
C	IJUP - Abstract	112
	Referências	114

Lista de Figuras

1.1	Diagrama do trabalho proposto.	5
2.1	Pontos característicos para reconhecimento a partir de impressão digital.	8
2.2	Exemplo de pontos característicos para reconhecimento a partir da retina.	8
2.3	Imagem de uma face e algumas medidas essenciais na fase do reconhecimento.	9
2.4	Imagem da disposição das veias na mão de um indivíduo recorrendo a radiação na região espectral dos infravermelhos.	10
2.5	Pontos característicos utilizados na geometria da mão e impressão palmar.	10
2.6	Dois tipos de equipamento utilizado na aquisição, câmara CCD e scanner de mesa.	12
2.7	Dois tipos de aquisição: a) com pinos fixadores (Ross et al., 1999); b) sem pinos fixadores.	13
2.8	Tipo de aquisição e características extraídas por Ross et al. (1999).	14
2.9	a) Pontos característicos da mão; b) Características utilizadas por Wong e Shi (2002).	17
2.10	Algoritmo de pré-processamento usado por Faundez-Zanuy e Merida (2005).	18
2.11	Método para detecção das características utilizado por Adan et al. (2008). a) Análise do contorno em coordenadas polares; b) Localização mais provável dos pontos detectados em (a) no contorno; c) Características utilizadas no reconhecimento.	19
2.12	Método para detecção dos pontos característicos baseado na análise da curvatura do contorno (Montalvao et al., 2010).	20
2.13	Método proposto por Montalvao et al. (2010) e suas fases constituintes.	20
3.1	Scanner utilizado na fase de aquisição.	23
3.2	Alguns exemplos de imagens adquiridas (de notar a liberdade na posição, orientação e configuração da mão).	24
4.1	Sistema de coordenadas de referência usado.	27
4.2	Exemplo de binarização a partir da definição de um valor de <i>threshold</i>	28
4.3	Espaço de cores HSV (http://clubelek.insa-lyon.fr/ressources/video/index2.php).	29
4.4	a) Imagem original no espaço de cores RGB, b) Representação do canal <i>Hue</i> , em escala de cinzento, da imagem a).	30
4.5	a) e b) Histogramas representativos da distribuição dos valores do matiz em duas amostras distintas.	31
4.6	a) Imagem original no espaço de cores RGB, b) Primeiro ensaio da obtenção da máscara binária correspondente à imagem em a).	32
4.7	a) Imagem original no espaço de cores RGB, b) Segundo ensaio da obtenção da máscara binária correspondente à imagem em a), agora com aplicação de alguns filtros.	33

4.8	a) Atribuição de uma etiqueta diferente a cada objecto separado dos restantes; b) Detecção do objecto com maior área e eliminação dos restantes - Máscara binária final.	34
4.9	a) Imagem original com dedos juntos; b) Respectiva máscara binária com falha na separação de todos os dedos.	35
4.10	a) Imagem original com tonalidade mais rosada; b) Respectiva máscara binária mostrando deficiências nas extremidades da mão e as "cristas" nos vales.	35
4.11	a) Imagem original com todos os dedos completamente juntos; b) Respectiva máscara binária definindo claramente a zona entre os dedos.	37
4.12	a) Imagem original com alguns dedos juntos; b) Respectiva máscara binária definindo claramente a zona entre os dedos.	38
4.13	a) Imagem original com todos os dedos abertos e uma ligeira sombra nos dedos indicador e médio; b) Respectiva máscara binária (de notar que existe mais que um objecto); c) Máscara binária após etiquetagem dos vários componentes.	38
4.14	a) Imagem original de indivíduo de raça negra com todos os dedos afastados; b) Respectiva máscara binária (de notar que existe mais que um objecto); c) Máscara binária após etiquetagem dos vários componentes.	39
4.15	a) Imagem original RGB; b) Imagem original em tons de cinzento; c) Imagem original em tons de cinzento, resultado da aplicação de <i>imadjust()</i>	41
4.16	a) Imagem original RGB; b) Respectiva máscara binária mostrando os bons resultados do método de <i>Otsu</i>	41
4.17	a) Imagem original RGB com presença de sombras; b) Respectiva máscara binária mostrando os bons resultados do método de <i>Otsu</i> para este caso.	42
4.18	a) Imagem original RGB com dedos juntos; b) Respectiva máscara binária mostrando os bons resultados do método de <i>Otsu</i> para este caso.	42
4.19	a) Imagem original RGB com dedos completamente juntos; b) Respectiva máscara binária mostrando a incapacidade de uma correcta segmentação da mão usando o método de <i>Otsu</i>	43
4.20	a) Imagem original RGB de individuo de raça negra; b) Respectiva máscara binária usando o método de <i>Otsu</i> , omitindo algumas zonas da mão devido a uma má segmentação.	43
4.21	a) Imagem original RGB com dedos completamente juntos; b) Respectiva máscara binária mantendo conectividade entre todas as zonas da mão e, ao mesmo tempo,garantindo a separação completa entre os dedos. <i>Threshold</i> aplicado: 117/255.	45
4.22	a) Imagem original RGB de individuo de raça negra; b) Respectiva máscara binária mantendo conectividade entre todas as zonas da mão e, ao mesmo tempo,garantindo a separação completa entre os dedos. <i>Threshold</i> aplicado: 53/255.	45
4.23	a) Imagem original RGB com dedos completamente juntos; b) Respectiva máscara binária mantendo conectividade entre todas as zonas da mão mas, ao mesmo tempo, não separando devidamente os dedos. <i>Threshold</i> aplicado: 122/255; c) Respectiva máscara binária garantindo uma correcta separação entre dedos mas não assegurando total conectividade entre todas as zonas da mão. <i>Threshold</i> aplicado: 132/255.	46

4.24	a) Imagem original RGB com dedos completamente juntos; b) Respectiva máscara binária mantendo conectividade entre todas as zonas da mão mas, ao mesmo tempo, não separando devidamente os dedos. <i>Threshold</i> aplicado: 118/255; c) Respectiva máscara binária garantindo uma correcta separação entre dedos mas não assegurando total conectividade entre todas as zonas da mão. <i>Threshold</i> aplicado: 155/255.	46
4.25	a) Imagem original RGB; b) Respectiva máscara binária usando o método de Otsu; c) Contorno da mão usando o método de <i>Moore Neighbor</i>	49
4.26	a) Imagem original RGB; b) Respectiva máscara binária usando o método de Otsu; c) Contorno da mão usando o método de <i>Moore Neighbor</i>	49
4.27	a) Imagem original em tons de cinzento; b) Respectiva imagem gradiente, resultado da aplicação do algoritmo de <i>Canny</i>	51
4.28	a) Imagem gradiente de <i>Canny</i> depois de aplicado o <i>threshold</i> para manter apenas as bordas mais fortes; b) Respectivo esqueleto (resultado da aplicação da operação morfológica de esqueletização); c) Esqueleto final, após aplicação da operação morfológica de poda.	52
4.29	Método de alinhamento proposto por Wong e Shi (2002) e Adan et al. (2008). . .	54
4.30	Sistemas de referência do objecto e final.	56
4.31	a) Imagem original RGB; b) Respectiva imagem do contorno no sistema referencial inicial com o CM assinalado. Apresenta um ângulo de inclinação: $\theta = 75.06^\circ$, 1ºQuadrante; c) Respectiva imagem do contorno alinhada com o sistema referencial final.	58
4.32	a) Imagem original RGB com a mão invertida; b) Respectiva imagem do contorno no sistema referencial inicial com o CM assinalado. Apresenta um ângulo de inclinação: $\theta = -46.53^\circ$, 4ºQuadrante; c) Respectiva máscara binária alinhada com o sistema referencial final.	58
4.33	a) Imagem original RGB com a mão disposta na horizontal; b) Respectiva imagem do contorno no sistema referencial inicial com o CM assinalado. Apresenta um ângulo de inclinação: $\theta = 10.95^\circ$, 1ºQuadrante; c) Respectiva imagem do contorno alinhada com o sistema referencial final.	59
4.34	Passagem de uma imagem pela fase de pré-processamento.	60
5.1	Representação dos 9 pontos característicos.	62
5.2	a) Imagem original alinhada do contorno; b) Imagem do contorno original cortada ao nível do CM.	62
5.3	Imagem do contorno da parte superior da mão em coordenadas polares.	63
5.4	a) Pontos do contorno mapeados num espaço cartesiano; b) Comparação entre contorno original e contorno suavizado (com método <i>loess</i>), linha a vermelho - contorno original, linha a azul - contorno suavizado. Representação dos pontos identificados como os de maior curvatura, resultado da análise da primeira derivada.	64
5.5	Mapeamento dos pontos identificados na imagem original comprovando os bons resultados que este método apresenta.	64
5.6	a) Imagem original alinhada do contorno; b) Imagem do contorno original cortada ao nível horizontal do CM. Neste caso apenas 4 dedos se encontram na imagem.	65
5.7	Imagem do contorno da parte superior da mão em coordenadas polares (4 dedos apenas).	66
5.8	a) Máscara binária alinhada da mão; b) Máscara binária com representação do respectivo esqueleto; c) Extracção do esqueleto da mão para futuro tratamento.	66

5.9	Método para detecção dos pontos correspondentes às pontas dos dedos a partir das pontas do esqueleto (pontos marcados a vermelho correspondem às pontas do esqueleto identificadas antes).	68
5.10	Esqueleto da máscara binária da mão invertido para detecção dos pontos correspondentes aos vales entre os dedos.	69
5.11	Método para detecção dos pontos correspondentes aos vales entre os dedos a partir de varrimentos entre linhas que percorrem os dedos.	70
5.12	a) Imagem original RGB; b) 1º passo do método: seguimento do contorno entre ponta do dedo mínimo e do dedo anelar (contorno original a vermelho, contorno suavizado a azul); c) 2º passo do método: seguimento do contorno entre as pontas dos dedos anelar e médio (contorno original a vermelho, contorno suavizado a azul); d) Resultado final do método.	71
5.13	a) Esqueleto apresentando pequenos segmentos que permaneceram após a operação de poda e que influenciam a detecção de pontos (pontas marcadas a vermelho correspondem a pontos verdadeiros, ponta marcada a laranja corresponde a ponto marcado erradamente); b) Esqueleto de uma mão em que os dedos se encontram muito juntos, apresentando a junção de dois segmentos em que uma ponta verdadeira é eliminada.	73
5.14	a) e b) Área convexa que envolve a mão (a vermelho), resultado da aplicação de <i>convexhull()</i> e o respectivo contorno (a azul).	74
5.15	a) Marcação dos pontos correspondentes aos vértices da área convexa; b) e c) Identificação dos pontos dos vértices da área convexa que coincidem com as pontas dos dedos.	75
5.16	Exemplos de imagens que permaneceram na BD depois de reduzida e retiradas aquelas em que a mão se encontrava com os dedos juntos.	76
6.1	Ilustração do método utilizado para detecção dos dois pontos da extremidade da mão correspondentes às bases esquerda do dedo mínimo e base direita do dedo indicador. Pontos a azul representam o trecho do contorno percorrido, pontos marcados a vermelho representam o ponto existente para o qual é calculada a distância aos pontos percorridos e os pontos marcados a amarelo representam os novos pontos identificados.	81
6.2	Ilustração do método utilizado para extracção dos comprimentos $L_{1,4}$	82
6.3	Ilustração do método utilizado para determinação das larguras de cada dedo.	83
6.4	Ilustração do método utilizado para delimitação da região da ponta de cada dedo.	84
7.1	Gráfico representativo do comportamento da taxa de eficácia obtida pelo método de <i>feature matching</i> introduzindo um atributo de cada vez, pela ordem obtida através das correlações.	94
7.2	<i>Feature matching</i> - Gráfico representativo da evolução da taxa de eficácia obtida na experiência presente na Tabela 7.2.	95
7.3	FAR vs FRR. Curva a Vermelho - FRR; Curva a azul - FAR; Ponto marcado a amarelo - EER.	96
7.4	Curva de funcionamento do sistema ROC - Receiver Operator Characteristic. Ponto A: funcionamento para aplicações de elevada segurança; Ponto B: funcionamento para aplicações de baixa segurança.	97
7.5	SVM - Gráfico representativo da evolução da taxa de eficácia obtida na experiência presente na Tabela 7.5.	99
7.6	Exemplo da estrutura básica de uma árvore de decisão.	100

A.1	EPSON PERFECTION V300 PHOTO.	108
A.2	8 examples of images in the database. Please notice the freedom given to the individual on the hand positioning and finger arrangement.	108
B.1	Acquisition with peg's.	109
B.2	Examples of restrictions avoided in our Database. a) big part of the arm in image; b) orientation free; c) hand closed.	110
B.3	Characteristic Points to be detected.	110
B.4	Some samples of images in our Database.	111

Lista de Tabelas

2.1	Comparação feita por Jain et al. (2000) das diferentes tecnologias biométricas quanto aos requisitos.	11
2.2	Resultados da avaliação na identificação para métodos individual e de fusão (Tsapatsoulis e Pattichis, 2009). EER - Equal Error Rate; IE - Identification Error (1-GAR).	16
4.1	Diferentes casos no cálculo da orientação.	56
5.1	Taxas de eficácia na detecção dos pontos característicos da mão.	75
6.1	Excerto do conjunto de atributos extraídos para a fase de reconhecimento (as cinco imagens pertencem à mesma pessoa).	85
6.2	Características estatísticas dos atributos extraídos.	86
6.3	Resultados da medida de correlação de <i>Pearson</i> para: Atributos vs Classes. Atributos a verde: bem correlacionados; Atributos a vermelho: correlacionados negativamente; Atributo a cinzento: bem correlacionado mas devido à sua elevada escala infere instabilidade na classificação.	87
6.4	Resultados da medida de correlação de <i>Pearson</i> entre os quatro dedos.	87
6.5	Resultados da medida de correlação de <i>Pearson</i> entre comprimento e largura de cada dedo.	88
6.6	Resultados da medida de correlação de <i>Pearson</i> entre comprimento e área de cada dedo.	88
6.7	Lista dos atributos ordenada pela importância dos mesmos, utilizando dois métodos distintos.	89
7.1	Resultados para 50 experiências variando, em todas elas, os conjunto de treino e teste e a ordem de inserção dos atributos - <i>feature matching</i>	93
7.2	Resultados para uma das 50 experiências utilizando conjunto aleatório com ordem de inserção determinada pela análise das correlações - <i>feature matching</i> . Para esta experiência teve-se: 110 imagens para treino e 77 imagens para teste.	94
7.3	Resultados da GAR e FAR para os diferentes valores de <i>threshold</i> - <i>feature matching</i>	96
7.4	Resultados para 50 experiências variando, em todas elas, os conjunto de treino e teste e a ordem de inserção dos atributos - SVM.	98
7.5	Resultados para uma das 50 experiências utilizando conjunto fixo com ordem de inserção determinada pela análise das correlações - SVM.	99
7.6	Resultados de classificação utilizando árvores de decisão.	101

Abreviaturas e Símbolos

2D	Bidimensional
3D	Tridimensional
BD	Base de Dados
CCD	<i>Charge Coupled Device</i>
CM	Centro de Massa
DMA	Distância Média Absoluta
EER	<i>Equal Error Rate</i>
EQM	Erro Quadrático Médio
FAR	<i>False Acceptance Rate</i>
FEUP	Faculdade de Engenharia da Universidade do Porto
FRR	<i>False Rejection Rate</i>
GAR	<i>Genuine Acceptance Rate</i>
HSL	Espaço de cores (<i>Hue, Saturation, Lightness</i>)
HSV	Espaço de cores (<i>Hue, Saturation, Value</i>)
ICIAR	<i>International Conference on Image Analysis and Recognition</i>
IJUP	Investigação Jovem da Universidade do Porto
INESC	Instituto de Engenharia de Sistemas e Computadores
KNN	<i>K-Nearest Neighbour Classifier</i>
MIEEC	Mestrado Integrado em Engenharia Electrotécnica e de Computadores
MLE	<i>Maximum Likelihood Estimative</i>
RBF	<i>Radial Basis Function</i>
RGB	Espaço de cores (<i>Red, Green, Blue</i>)
RN	Rede Neuronal
ROC	<i>Receiver Operator Characteristic</i>
SVM	<i>Support Vector Machine</i>
UTM	Unidade de Telecomunicações e Multimédia

Capítulo 1

Introdução

1.1 Enquadramento

Este trabalho de dissertação surge no âmbito do Mestrado Integrado em Engenharia Electrotécnica e de Computadores (MIEEC), major de Automação, da Faculdade de Engenharia da Universidade do Porto (FEUP). Foi realizado em ambiente empresarial no Instituto de Engenharia de Sistemas e Computadores do Porto (INESC Porto), na Unidade de Telecomunicações e Multimédia (UTM) com a supervisão do professor doutor Jaime dos Santos Cardoso. Este trabalho encontra-se também inserido no grupo Bio Star. Este grupo é constituído por estudantes, dos vários ciclos do ensino superior, das áreas da engenharia e da bioengenharia e tem como objectivo o desenvolvimento tecnológico nas áreas ”*bio-relacionadas*”¹.

1.2 Motivação

Os sistemas de reconhecimento biométrico são uma ferramenta com uma importância crucial no mundo da segurança e com uma crescente utilização em todo o tipo de serviços que implique a autenticação de pessoas. Devido à sua grande evolução nos últimos anos, estes sistemas têm vindo a tornar-se cada vez mais simples, baratos e adaptáveis a qualquer tipo de aplicação como é o exemplo de sistemas de reconhecimento através de impressões digitais (mais comuns), através da íris (mais sofisticado mas também mais dispendioso e intrusivo) ou através da face. No entanto, sistemas de reconhecimento baseados na mão, quer na geometria da mão quer na impressão da palma, estão ainda longe do favoritismo que outros tipos de sistemas já possuem. Esta impopularidade deve-se essencialmente a dois factores: à pouca facilidade de utilização e fraca simplicidade de interacção com o utilizador dos sistemas; e ao baixo desempenho que estes sistemas apresentam devido às grandes dimensões que as BD atingem.

¹<http://biostar.fe.up.pt/>

Pretende-se neste trabalho superar ao máximo as limitações dos sistemas de reconhecimento baseados na mão, ainda que, aqui será apenas abordada a vertente que utiliza a geometria da mão como informação biométrica para o reconhecimento.

Os sistemas actuais apresentam uma grande limitação a nível da Base de Dados (BD). Devido à elevada dimensão das imagens a tratar e dado serem utilizadas, normalmente, cerca de uma dezena de imagens de cada indivíduo no sistema (uma metade para treino e metade para teste) (Faundez-Zanuy e Merida, 2005; Ross et al., 1999), as BD rapidamente atingem dimensões críticas o que afecta o desempenho do sistema de reconhecimento.

A outra grande restrição está na aquisição de imagens para o reconhecimento/identificação. A aquisição de imagens da mão pode ser feita através de uma scanner ou de uma câmara, no entanto existem bastantes limitações quanto ao posicionamento da mão aquando da aquisição, por exemplo, a mão deve apresentar sempre uma determinada posição e orientação, a configuração dos dedos é imposta, etc. Em alguns sistemas são utilizados fixadores para obrigar o utilizador a colocar a mão na posição determinada, no entanto, para além de desconfortáveis estes fixadores causam deformações na mão que, posteriormente, irão influenciar a extracção de características. Outros sistemas utilizam já uma metodologia *peg's-free*, ou seja, é feita uma aquisição sem fixadores, o que permite um pouco mais de liberdade ao utilizador, mas, ao mesmo tempo, é-lhe imposta uma posição pré-determinada e caso o utilizador não obedeça a essa posição, o reconhecimento irá falhar.

A grande motivação para a realização deste trabalho passa então por tentar desenvolver um sistema capaz de mostrar um desempenho aceitável para universos maiores e também por eliminar todas ou, pelo menos, grande parte das restrições impostas na aquisição. Seria uma grande inovação a implementação de um sistema de reconhecimento biométrico através da geometria da mão, onde não houvesse quaisquer restrições no posicionamento da mão, onde esta pudesse estar rodada, invertida, com os dedos juntos ou afastados, etc..

1.3 Objectivos

Dada a motivação para este trabalho os objectivos surgem naturalmente:

1. Compreender em que consiste um sistema de reconhecimento biométrico:
 - Adquirir conhecimento sobre biometria, mais especificamente a biometria da mão;
 - Conhecer o modo de funcionamento de um sistema de reconhecimento biométrico e as técnicas utilizadas no mesmo;
2. Criação de uma Base de Dados de imagens da mão;
3. Inovar e/ou melhorar as técnicas já existentes, através da eliminação das restrições impostas pelos sistemas actuais relativamente ao posicionamento da mão aquando da aquisição:

- Levantamento de informação sobre esta matéria (revisão da literatura) e análise do máximo de técnicas possível para posterior comparação entre as mesmas;
- Implementação e teste das técnicas mais convenientes e comparação dos resultados obtidos por cada uma delas;
- Implementação de novas técnicas e comparação dos resultados com os das técnicas já existentes;

Nos dias de hoje, com toda a evolução associada a estes sistemas, todo o tipo de sistemas tecnológicos se querem o mais simples e fácil de utilizar. Quando falamos em sistemas de reconhecimento/identificação, este conceito toma dimensões maiores. Estes sistemas são implementados na entrada de edifícios, empresas ou espaços e devem tomar apenas poucos segundos ao utilizador. Assim, qualquer restrição que influencie o tempo ou facilidade de utilização é altamente penalizada pelos compradores ou potenciais interessados nestes sistemas.

A exclusão das restrições aquando da aquisição permitirá um avanço significativo destes sistemas, principalmente porque fará deles tecnologias mais simples de utilizar e com uma melhoria considerável na interacção utilizador/máquina. Um sistema em que o utilizador possa colocar a mão e ser identificado, sem qualquer cuidado na posição, terá inúmeras vantagens em comparação com um em que o utilizador tenha de ter em atenção a posição da mão, a colocação dos dedos nos fixadores/marcas e tenha, por vezes, de fazer várias tentativas de identificação para ser permitido o seu acesso.

Quando nos comprometemos a atingir todos estes objectivos torna-se fundamental a criação de uma Base de Dados. As BD já existentes, além de escassas, são bastante limitadas pois contêm apenas imagens de sistemas em que são utilizados fixadores ou onde existem todas as restrições anteriormente mencionadas. Assim, é um objectivo importante deste trabalho a criação de uma BD, em condições naturais de iluminação, com imagens adquiridas com total liberdade no posicionamento, orientação e configuração por parte do utilizador.

1.4 Descrição do Trabalho Proposto e Metodologia Utilizada

Primeiramente, era decisivo que fosse feito um estudo exaustivo do estado da arte relacionado com o tema em causa. Só assim se conseguia ter uma boa e correcta percepção de como funcionam estes sistemas e quais as características essenciais para obtenção de bons resultados. Neste estudo foram também investigadas algumas das metodologias já implementadas, de forma a melhorar alguns aspectos das mesmas, ou até utilizar outras técnicas já existentes com aplicabilidade diferente, como técnicas de binarização ou de análise de curvatura.

Após este estudo seguem-se quatro grandes fases: o pré-processamento, a detecção de pontos, a extracção de características e o reconhecimento.

Durante a fase de pré-processamento serão aplicados alguns algoritmos de modo a melhorar a qualidade da imagem, com vista à sua utilização nos processos seguintes. Esta fase engloba os passos de binarização (obtenção da máscara binária da mão), cálculo do centro de massa, orientação de mão e, por fim, a extracção de contornos.

Na fase de detecção de pontos serão implementadas várias técnicas para extracção dos pontos característicos da mão. Os pontos a serem extraídos nesta fase correspondem às cinco pontas dos dedos e aos quatro vales entre dedos.

A fase de extracção de características trata de, a partir dos pontos detectados anteriormente, calcular determinadas distâncias entre os pontos característicos que mais tarde serão utilizadas no reconhecimento.

Por fim, na fase de reconhecimento será feita a identificação da pessoa. Reconhecimento significando a identificação cognitiva de uma pessoa.

Neste trabalho foram então abordadas duas grandes temáticas, o processamento e análise de imagem onde serão estudadas e aplicadas técnicas de processamento de imagem, e *Machine Learning* onde foram implementados alguns dos algoritmos de aprendizagem mais comuns nesta área.

Revisão Tecnológica

O sistema a ser desenvolvido engloba duas grandes áreas:

- Processamento e análise de imagem, onde será utilizado o software MatLab devido ao seu enorme potencial em termos de funções e *toolboxes* específicas para processamento e análise de imagem.
- *Machine Learning* (Aprendizagem Automática), que será também implementado usando o software MatLab, pois este incorpora praticamente todos os algoritmos de aprendizagem necessários à realização deste trabalho. Será também utilizado o software WEKA 3-4 para fins de comparação de resultados com os restantes classificadores. O WEKA é uma ferramenta de código aberto, implementado em *Java*, que incorpora inúmeros algoritmos de aprendizagem de dados convencionais.

O sistema apresenta os seguintes blocos principais:

1. Aquisição de imagens: construção da BD.
2. Pré-processamento
 - (a) Binarização
 - (b) Extracção de Contornos
 - (c) Alinhamento da Imagem

3. Extracção de Características
 - (a) Detecção dos Pontos Característicos
 - (b) Medição das Características a usar no Reconhecimento
4. Fase de anotação da BD: Construção da Base de Dados com as Assinaturas dos Utilizadores
5. Fase de Identificação: Recurso a Classificadores ou Funções de Matching para a Identificação do Utilizador.

A Figura 1.1 mostra um diagrama do trabalho proposto, para uma melhor compreensão do mesmo.

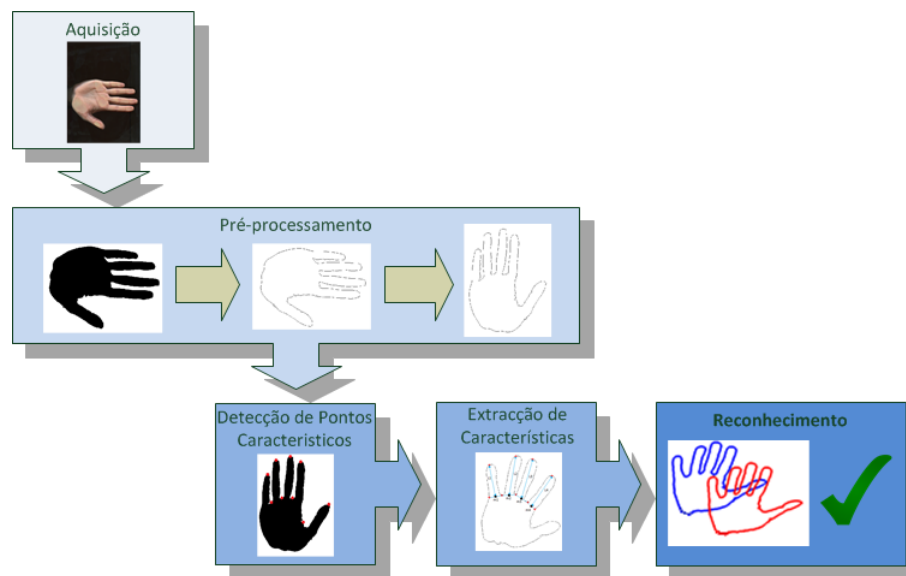


Figura 1.1: Diagrama do trabalho proposto.

1.5 Contribuições

- Foi desenvolvido um novo algoritmo que utiliza o esqueleto da imagem binária para detecção dos pontos correspondentes às pontas dos cinco dedos. Este método, para detecção dos três vales entre dedos, utilizava um algoritmo de seguimento de contornos já suavizado. Como resultado, apresentou uma taxa de 97% de acerto nos pontos pretendidos, sendo que foi aplicado em imagens de uma BD sem restrições na aquisição. Este facto constituiu uma grande contribuição para esta matéria.
- Todos os resultados obtidos durante a fase de teste constituem também uma enorme contribuição para este trabalho. Ainda que a grande parte dos métodos/algoritmos testados não tenha surtido o resultado esperado, o seu teste foi de extrema importância, pois mesmo não

apresentando resultados positivos constitui uma enorme e valiosa fonte de informação para futuros investigadores nesta área.

- Criação de uma nova BD com imagens adquiridas sem quaisquer restrições quanto ao posicionamento da mão. Além de esta ser uma BD sem antecedentes, influenciou numa maior motivação para o estudo deste tipo de sistemas e respectivas restrições, dada ser uma BD pública e de utilização gratuita.
- Foi também submetido um *abstract* para apresentação oral na conferência IJUP2011 - Investigação Jovem da Universidade do Porto.
- Por fim, foi organizada uma competição internacional de algoritmos "*Hand Geometric Points Detection Competition*"² para detecção dos pontos característicos da mão. Esta competição está inserida na conferência internacional **ICIAR 2011** - *International Conference on Image Analysis and Recognition*³, a realizar-se no Canadá e dará hipótese a que os participantes com melhores resultados apresentem um artigo para submissão nesta conferência.

1.6 Estrutura do Documento

No capítulo 2, é exposta toda a informação recolhida no estudo do estado da arte. Aí é explicada a arquitectura e o funcionamento de um sistema típico de reconhecimento biométrico. Serão também identificados e expostos os métodos e algoritmos mais relevantes nesta matéria.

No capítulo 3 é explicado o método aqui adoptado para aquisição das imagens. Será também mencionada a BD criada e as suas principais características.

O capítulo 4 descreve as metodologias utilizadas para o pré-processamento de imagem, onde inclui todo um conjunto de algoritmos para tratamento da imagem. Neste capítulo serão explicados os métodos utilizados no *enhancement* da imagem de modo a alcançar os requisitos necessários à fase seguinte.

Após a fase de pré-processamento, no capítulo 5, serão demonstrados todos os métodos utilizados para a extracção dos pontos característicos, os que produzem bons resultados e também todos aqueles que foram testados e não apresentaram resultados satisfatórios.

No capítulo 6 serão apresentados os métodos utilizados para extracção e selecção de características a serem utilizadas nas fases seguintes.

O capítulo 7 apresenta a fase final do sistema, o reconhecimento. Aqui são apresentados os vários métodos de classificação utilizados para a identificação de pessoas no sistema e os respectivos resultados obtidos.

O capítulo 8 finaliza este relatório apresentando os resultados e conclusões gerais do sistema desenvolvido e respectivos passos intermédios. São também aqui expostos alguns tópicos de trabalho a desenvolver-se futuramente.

²<http://paginas.fe.up.pt/hgc2011/>

³<http://www.iciar.uwaterloo.ca/iciar11/>

Capítulo 2

Reconhecimento Biométrico

2.1 Sistemas Biométricos

A biometria é o estudo das características físicas ou comportamentais dos humanos. Recentemente, nas últimas décadas, este termo tem sido também alvo de estudo como forma de identificação de pessoas, sendo usado em diversas aplicações, tais como na identificação criminal e controlo de acessos, entre outras. A biometria, no âmbito do reconhecimento, recorre a partes específicas do corpo humano para a extracção de características a serem usadas no reconhecimento, como a palma da mão, os olhos, as impressões digitais, a retina, a íris, a geometria da mão, entre outros (Jain e Ross, 2004). A premissa em que estes sistemas se fundamentam é a de que cada indivíduo é único e possui características físicas distintas que permitem a sua distinção de entre um determinado grupo de indivíduos (Prabhakar et al., 2003). Um sistema de reconhecimento biométrico, dependendo da região anatómica a que está associado, extrai as características mais relevantes dessa parte do corpo e com esse conjunto de características é criada uma "assinatura" da pessoa correspondente. Esta "assinatura" representa a identidade de determinada pessoa no sistema e posteriormente, aquando do reconhecimento, é comparada com o conjunto de características inserido e, dependendo da medida de similaridade, a pessoa é correctamente identificada ou não, podendo em certos casos ser recusada a sua identificação. De entre os vários tipos de reconhecimento biométrico existem uns que se destacam pela sua elevada eficácia e segurança, tais como as impressões digitais ou o reconhecimento da íris ou da retina, e existem outros que se salientam pelo seu carácter *user-friendly*, simplicidade e baixo custo, como é o caso do reconhecimento facial ou através da mão. No trabalho de Kounoudes et al. (2008), pode assistir-se à implementação de alguns dos métodos de reconhecimento e também a uma fusão das saídas de cada um deles e perceber assim o resultado de um sistema biométrico multi-modal.

O reconhecimento a partir da impressão digital de uma pessoa é, actualmente, um dos métodos de reconhecimento mais popular em todo o mundo. Apresenta uma elevada eficácia nas taxas de identificação e os sistemas são cada vez mais simples e fáceis de utilizar. Na Figura 2.1 pode ser observada a identificação de alguns pontos característicos das impressões digitais. No entanto, estes sistemas são, por vezes, considerados um pouco intrusivos na medida em que utilizam infor-

mação privada da pessoa (parte do seu registo civil) e a guardam nas suas bases de dados e são também um pouco dispendiosos para determinadas aplicações. Uma desvantagem deste método em comparação com os baseados na geometria da mão reside no equipamento que cada um utiliza, pois para a geometria da mão um *scanner* comercial pode ser utilizado, enquanto para a aquisição de impressões digitais são necessários equipamentos mais sofisticados.



Figura 2.1: Pontos característicos para reconhecimento a partir de impressão digital (http://www.filebuzz.com/files/Biometric_Recognition/).

Quando se trata de projectar um sistema de alta segurança, em que é preferível garantir uma taxa de intrusão (*False Acceptance Rate* - FAR) muito próxima de zero mesmo que isso implique um decréscimo da eficácia (*Genuine Acceptance Rate* - GAR), então os sistemas baseados na íris ou retina dos olhos são os preferidos (ver Figura 2.2). Apesar de ser uma parte do corpo visível, é uma componente interna do olho e a probabilidade de existirem duas íris iguais estima-se que seja cerca de $1/(10^{78})$. Dado o carácter intrusivo da aquisição deste tipo de sistemas, estes são utilizados apenas em sistemas de muito alta segurança.

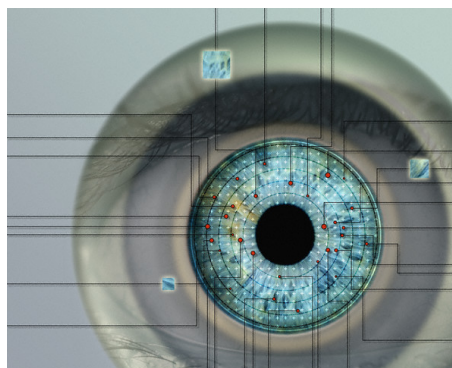


Figura 2.2: Exemplo de pontos característicos para reconhecimento a partir da retina (<http://www.biometricvisions.com/technology/>).

Quando se pretende a adopção de um sistema cuja aquisição seja menos intrusiva, o reconhecimento baseado em imagens da face é uma solução viável. Estes sistemas baseiam-se na aquisição de uma imagem facial e através de determinadas características extraídas é feita uma comparação com as imagens existentes na BD. As características mais relevantes para o reconhecimento são as distancias entre olhos, nariz, queixo e linha dos cabelos (ver Figura 2.3). O reconhecimento facial possui ainda algumas limitações, nomeadamente ao nível das imagens que captura, pois pode ser confundido, por exemplo, através das expressões faciais, pêlos e óculos, dificultando a identificação do utilizador. Contudo, com o desenvolvimento desta tecnologia em 3 D, muitas dessas dificuldades deixaram de existir, mas ao mesmo tempo tornam-se sistemas mais caros e, portanto, com menor atracção para as aplicações mais comuns.

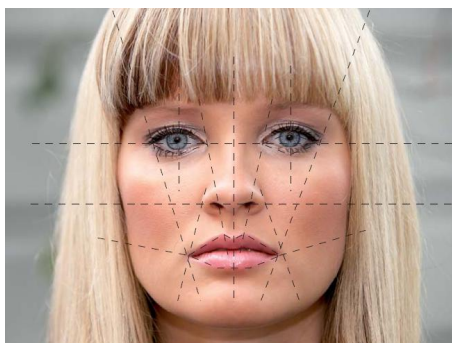


Figura 2.3: Imagem de uma face e algumas medidas essenciais na fase do reconhecimento (<http://www.biometry.com/>).

Com o desenvolvimento da biometria associada ao reconhecimento, novas técnicas e novos tipos de sistemas começam a ser investigados e desenvolvidos, como é o caso do reconhecimento das veias. Estes sistemas são mais recentes e fundamentam-se na disposição das veias, tal como é ilustrado na Figura 2.4, que se tem como distinta de indivíduo para indivíduo. São sistemas que se encontram numa gama de custo médio, devido principalmente ao tipo de equipamento utilizado aquando da aquisição, mas que acarretam ao mesmo tempo, uma segurança considerável, pois o padrão das veias é diferente mesmo entre gémeos. O facto de algumas veias não serem visíveis a olho nu através da pele faz com que a falsificação e manipulação destas sejam praticamente nulas.

Por fim, um tipo de sistema de reconhecimento bastante popular é o baseado na identificação biométrica da mão. Este tipo de sistemas pode incidir sobre a geometria da mão, sobre a impressão palmar ou até sobre ambas (Kumar et al., 2003). A Figura 2.5 mostra uma imagem da mão em que estão representados os pontos característicos a serem usados na geometria e outra imagem de uma impressão palmar. A geometria da mão diz respeito a uma série de medidas que podem ser extraídas determinando alguns pontos característicos, tais como as pontas dos dedos e vales.



Figura 2.4: Imagem da disposição das veias na mão de um indivíduo recorrendo a radiação na região espectral dos infravermelhos (<http://v-newswire.com/technology/>).

Sistemas baseados na geometria da mão são normalmente aplicados em universos mais reduzidos, na ordem das centenas dado as suas características não serem completamente distintivas em universos maiores. Os sistemas que utilizam a impressão palmar como objecto de reconhecimento têm um funcionamento em algo semelhante aos sistemas de impressões digitais. Isto é, as características normalmente usadas para o reconhecimento neste tipo de sistemas são as linhas principais (linha da vida, da cabeça, do coração, etc), a textura da palma, ilhas e bifurcações (tal como nas impressões digitais), alguns pormenores únicos da palma, características estatísticas, entre muitas outras. Os sistemas baseados no reconhecimento da mão são muito atractivos, pois apresentam um baixo custo devido ao equipamento básico que utilizam na aquisição (um *scanner* básico comercial). No entanto, são ainda pouco simples e apresentam baixa facilidade de interacção por parte do utilizador.

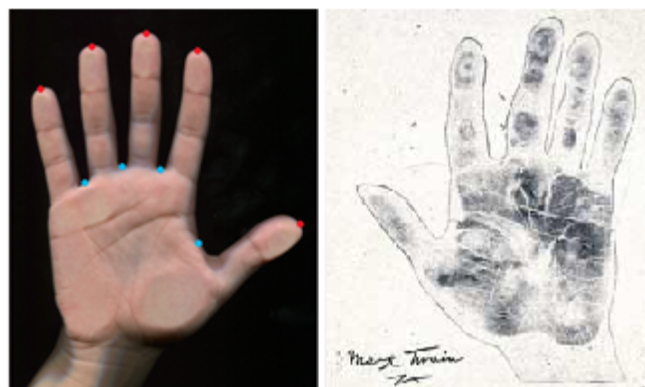


Figura 2.5: Pontos característicos utilizados na geometria da mão e impressão palmar (<http://www.nexistepas.com/2007/08/cheiros-language-of-the-hand/>).

A Tabela 2.1 apresenta um resumo das principais características a ter em conta nos sistemas de reconhecimento biométrico. De seguida são explicados os princípios tidos em conta para tal comparação (Jain et al., 2000):

- **Universalidade:** significa que todas as pessoas devem possuir a característica;
- **Singularidade:** indica que esta característica não pode ser igual em pessoas diferentes;
- **Permanência:** significa que a característica não deve variar com o tempo;
- **Mensurabilidade:** indica que a característica pode ser medida quantitativamente.
- **Desempenho:** refere-se à precisão de identificação, os recursos requeridos para conseguir uma precisão de identificação aceitável e ao trabalho ou factores ambientes que afectam a precisão da identificação;
- **Aceitabilidade:** indica o quanto as pessoas estão dispostas a aceitar os sistemas biométricos;
- **Protecção:** refere-se à facilidade/dificuldade de enganar o sistema com técnicas fraudulentas.

Tabela 2.1: Comparação feita por Jain et al. (2000) das diferentes tecnologias biométricas quanto aos requisitos.

Sistema Biométrico	Universalidade	Singularidade	Permanência	Mensurabilidade	Desempenho	Aceitabilidade	Protecção
Face	Alto	Baixo	Médio	Alto	Baixo	Alto	Baixo
Impressão Digital	Médio	Alto	Alto	Médio	Alto	Médio	Alto
Geometria da Mão	Médio	Médio	Médio	Alto	Médio	Médio	Médio
Veias	Médio	Médio	Médio	Médio	Médio	Médio	Alto
Íris	Alto	Alto	Alto	Médio	Alto	Baixo	Alto
Retina	Alto	Alto	Médio	Baixo	Alto	Baixo	Alto
Assinatura	Baixo	Baixo	Baixo	Alto	Baixo	Alto	Baixo
Voz	Médio	Baixo	Baixo	Médio	Baixo	Alto	Baixo

2.2 Reconhecimento Biométrico Baseado na Geometria da Mão

2.2.1 Modos de Aquisição

A primeira fase de um sistema de reconhecimento é a aquisição de imagens para a construção da BD. Normalmente são necessárias entre 5 a 15 imagens por pessoa para a implementação do sistema (Wong e Shi, 2002; Tsapatsoulis e Pattichis, 2009; Ross et al., 1999; Montalvao et al., 2010). Algumas são utilizadas para treinar o sistema e para criar a assinatura de cada indivíduo

representado na BD e as restantes são necessárias para teste do sistema, para assim se obter uma taxa de eficácia do mesmo.

Equipamentos de Aquisição Típicos

A aquisição das imagens para a base de dados é geralmente efectuada recorrendo a câmaras CCD (*charge coupled device*) (Ross et al., 1999), ou a scanners (Wong e Shi, 2002), estando ambos os equipamentos ilustrados na Figura 2.6. Normalmente a escolha do equipamento a ser utilizado depende da aplicação e também, essencialmente, da fase de pré-processamento e dos algoritmos aí utilizados, pois estes podem depender de algumas características tal como a resolução da imagem na aquisição. Por exemplo, se a finalidade da aplicação é o reconhecimento a partir da impressão da palma, imagens tiradas com uma câmara não serão as mais indicadas pois com este tipo de equipamento as linhas e textura da mão não são tão evidentes.



Figura 2.6: Dois tipos de equipamento utilizado na aquisição, câmara CCD e scanner de mesa (<http://www.digitalvix.com.br/> http://www.sbig.com/products/402_new.htm).

Métodos de Aquisição

Existem, basicamente, dois métodos para efectuar a aquisição de imagens para o reconhecimento baseado na mão. O primeiro passo essencial na fase subsequente do pré-processamento passa pela identificação da zona da mão na imagem e, para passar à fase seguinte, é necessário que a mão esteja numa determinada posição e orientação para que seja otimizada a extracção dos pontos característicos. Em Faundez-Zanuy e Merida (2005) podemos verificar alguns exemplos de imagens com aquisição defeituosa para este tipo de sistemas. Um dos métodos, o mais primitivo nesta matéria, utiliza um scanner para a aquisição onde existem uns fixadores que, tal

como pode ser constatado na Figura 2.7(a), obrigam o utilizador a colocar a mão numa determinada posição (dedos afastados) e orientação (vertical com dedos na parte superior), (Ross et al., 1999; Tsapatsoulis e Pattichis, 2009). Este método garante que, caso o utilizador obedeça à posição dos fixadores, a mão se encontre sempre na posição pré-definida e otimizada para a fase de pré-processamento. Pode ser observado em Struc e Pavesic (2009) um sistema de aquisição semelhante a este, no qual a plataforma de aquisição onde se deve colocar a mão contém uns fixadores com sensores de pressão. Assim, só quando o utilizador tiver a mão correctamente colocada na superfície, é que o sistema faz a aquisição a partir de uma câmara CCD. Outro método, mais recente, utiliza um sistema livre de fixadores, isto é, o utilizador tem uma maior liberdade, ainda que existam umas marcas que indicam a posição em que o utilizador deve colocar a mão, como acontece em Wong e Shi (2002); Montalvao et al. (2010). Neste método é já possível a aquisição tanto com scanner como com câmaras (ver Figura 2.7(b)).

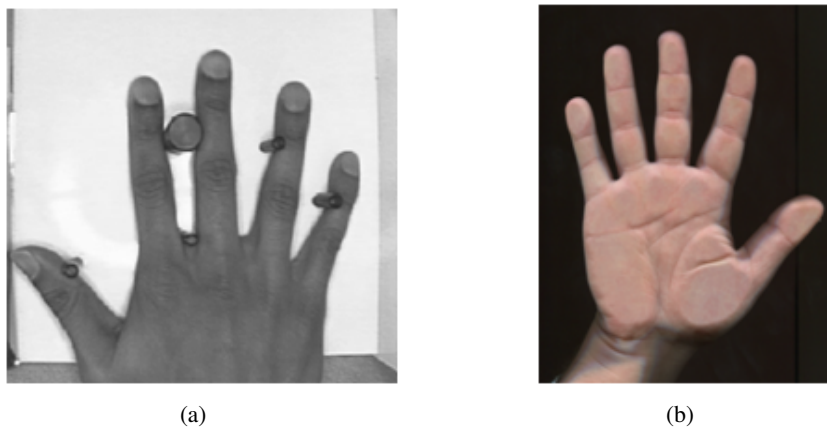


Figura 2.7: Dois tipos de aquisição: a) com pinos fixadores (Ross et al., 1999); b) sem pinos fixadores.

2.2.2 Detecção de Pontos, Extracção de Características e Resultados

Varrimentos na imagem em escala de cinzento

O reconhecimento baseado na geometria da mão existe já desde o início dos anos 70. No entanto, mesmo tendo sido alvo de muito estudo e investigação nos últimos anos, existe ainda pouca literatura sobre as técnicas mais usadas. Ross et al. (1999) apresentou um sistema de verificação

baseado na geometria da mão em que a aquisição é feita utilizando uma câmara CCD e cinco fixadores amovíveis para posicionamento da mão, (Figura 2.8). Este autor não apresenta um método de detecção de pontos característicos da mão e recorre directamente à extracção de características. Para isso utiliza um método baseado em varrimentos da imagem segundo uma determinada direcção onde se pretende que seja efectuada a medição, analisando a escala de cinzas da imagem original. O objectivo é determinar os dois pontos correspondentes ao início e fim do objecto que se pretende medir (dedo, palma, pulso, etc.). Estes pontos correspondem às duas transições abruptas existentes no perfil de intensidades para o respectivo varrimento. Para a detecção destes dois pontos, o autor utiliza dois métodos, um de estimação de parâmetros, onde recorre ao Estimador de Máxima Verosimilhança (MLE - *Maximum Likelihood Estimative*), e outro em que é utilizada uma janela com determinado comprimento que efectua um varrimento segundo a direcção determinada para detecção dos referidos pontos. Este sistema apresentou no final, para uma BD de 50 indivíduos uma taxa de eficácia de 85% com uma taxa de intrusão de 2%.

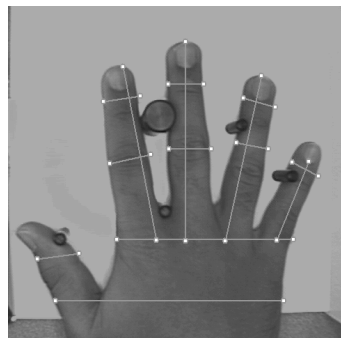


Figura 2.8: Tipo de aquisição e características extraídas por [Ross et al. \(1999\)](#).

Construção de *templates* baseados na silhueta e na máscara binária da mão, recorrendo a alguns descritores

[Tsapatsoulis e Pattichis \(2009\)](#) apresentaram um sistema de identificação e autenticação baseado na geometria da mão, com aquisição feita com recurso a fixadores e a uma câmara. Em primeiro lugar, são construídos *templates* baseados na silhueta e em imagens binarizadas da mão, recorrendo a quatro técnicas: os descritores de *Fourier*, o espectro de potência, medições da área da mão binarizada e o descritor visual **MPEG-7**. Posteriormente, foram investigadas abordagens para combinar os vectores de características. O sistema proposto foi testado na BD POLYBIO ([Antonakoudi A., 2009](#)) que consiste em 180 imagens de 45 indivíduos. Na fase de reconhecimento, o sistema foi testado num modo baseado em vectores de características e num outro baseado em

pontuação (obtida a partir das distâncias resultantes da comparação entre vectores de características), concluindo-se que o primeiro método apresenta melhores resultados tanto para identificação como para autenticação, exibindo um *Equal Error Rate* (EER) de 2.12% quando utiliza uma combinação entre a cor, a forma do contorno e o histograma do contorno. Alguns resultados obtidos por este autor estão representados na Tabela 2.2.

Análise do contorno da mão e respectiva curvatura

Com o avanço na investigação deste tipo de sistemas, começaram a aparecer sistemas de reconhecimento baseados na mão que fazem a aquisição de uma forma mais prática, isto é, não utilizam fixadores, tornando assim a fase da aquisição mais confortável e prática para o utilizador, como se verifica no trabalho apresentado por Wong e Shi (2002). Neste sistema é utilizado um sistema livre de fixadores que usa um scanner para a aquisição. Após esta fase a imagem é sujeita a uma etapa de pré-processamento, no sentido de segmentar a mão e obter o contorno da mesma. Posteriormente, é aplicado um algoritmo de seguimento de contornos que para cada ponto determina o valor da sua curvatura. A partir dos valores da curvatura é possível determinar os nove pontos característicos da mão (cinco pontas dos dedos e quatro vales entre eles - Figura 2.9(a)) e com eles calcular as características (Figura 2.9(b)) necessárias à identificação, que são:

- Comprimentos ($L_{1,4}$) dos quatro dedos: mínimo, anelar, médio e indicador;
- Larguras ($W_{1,8}$) dos mesmos quatro dedos em duas zonas específicas do dedo;
- Região ($S_{1,3}$) da ponta de três dedos.

Com base nas medidas alcançadas obtiveram 89% de eficácia e 2.2% de taxa de intrusão. Não obstante, é ainda referido pelo autor que existem algumas restrições no posicionamento da mão aquando da aquisição devido à necessidade de pinos fixadores. Este refere também que são apenas utilizadas medidas relativas aos quatro dedos dada a imprecisão das medidas obtidas para o dedo polegar.

De uma forma em tudo semelhante ao trabalho de Wong e Shi (2002) surgiu o trabalho realizado por Faundez-Zanuy e Merida (2005). O autor constrói uma base de dados com 22 indivíduos empregando o mesmo método de aquisição que em Wong e Shi (2002) e segue o mesmo modelo de sistema de reconhecimento, (ver Figura 2.10), recorrendo a uma fase de pré-processamento da imagem, onde é feita a binarização e conseqüente segmentação da mão. De seguida é também detectado o contorno da mão, utilizando aqui um filtro LoG, Laplaciano da função Gaussiana.

Tabela 2.2: Resultados da avaliação na identificação para métodos individual e de fusão (Tsapatsoulis e Pattichis, 2009). EER - Equal Error Rate; IE - Identification Error (1-GAR).

Method	Features	number of features	IE(%)	EER(%)
<i>FD</i>	<i>Fourier Descriptors</i>	8	20.00	14.15
<i>FD</i>	<i>Fourier Descriptors</i>	16	11.43	14.61
<i>FD</i>	<i>Fourier Descriptors</i>	32	12.14	16.62
<i>PS</i>	<i>Power Spectrum coefficients of contour</i>	8	20.71	14.18
<i>PS</i>	<i>Power Spectrum coefficients of contour</i>	16	13.57	13.84
<i>PS</i>	<i>Power Spectrum coefficients of contour</i>	24	15.00	16.10
<i>AF</i>	<i>Area related Features</i>	7	10.00	7.30
<i>CL</i>	<i>MPEG-7: Color layout descriptor</i>	12	5.71	6.29
<i>CS</i>	<i>MPEG-7: Contour shape descriptor</i>	5	43.57	15.48
<i>RS</i>	<i>MPEG-7: Region shape descriptor</i>	35	12.86	11.75
<i>EH</i>	<i>MPEG-7: Edge histogram descriptor</i>	80	10.71	7.16
<i>FF1</i>	<i>Concatenated FD and PS</i>	32	10.00	14.89
<i>FF2</i>	<i>Concatenated FD and AF</i>	23	6.43	9.93
<i>FF3</i>	<i>Concatenated PS and AF</i>	23	4.29	8.06
<i>FF4</i>	<i>Concatenated CL and CS</i>	17	1.43	3.87
<i>FF5</i>	<i>Concatenated CL and RS</i>	47	3.57	4.91
<i>FF6</i>	<i>Concatenated CL and EH</i>	92	3.57	3.33
<i>FF7</i>	<i>Concatenated CS and RS</i>	40	8.57	8.36
<i>FF8</i>	<i>Concatenated CS and EH</i>	85	5.00	5.76
<i>FF9</i>	<i>Concatenated RS and EH</i>	115	4.29	5.79
<i>FF10</i>	<i>Concatenated FD,PS and AF</i>	39	5.71	10.46
<i>FF11</i>	<i>Concatenated CL,CS and RS</i>	52	3.57	4.20
<i>FF12</i>	<i>Concatenated CL,CS and EH</i>	97	1.43	2.06
<i>FF13</i>	<i>Concatenated CL,RS and EH</i>	127	1.43	2.28
<i>FF14</i>	<i>Concatenated CS,RS and EH</i>	120	3.57	4.82
<i>SF1</i>	<i>Score Fusion of FD and PS</i>	32	9.29	14.73
<i>SF2</i>	<i>Score Fusion of FD and AF</i>	23	4.29	7.74
<i>SF3</i>	<i>Score Fusion of PS and AF</i>	23	4.29	6.81
<i>SF4</i>	<i>Score Fusion of CL and CS</i>	17	0.71	4.16
<i>SF5</i>	<i>Score Fusion of CL and RS</i>	47	0.00	4.55
<i>SF6</i>	<i>Score Fusion of CL and EH</i>	92	1.43	2.85
<i>SF7</i>	<i>Score Fusion of CS and RS</i>	40	9.29	7.78
<i>SF8</i>	<i>Score Fusion of CS and EH</i>	85	5.00	5.32
<i>SF9</i>	<i>Score Fusion of RS and EH</i>	115	4.29	5.09
<i>SF10</i>	<i>Score Fusion of FD,PS and AF</i>	39	2.86	7.30
<i>SF11</i>	<i>Score Fusion of CL,CS and RS</i>	52	0.00	2.73
<i>SF12</i>	<i>Score Fusion of CL,CS and EH</i>	97	0.71	2.12
<i>SF13</i>	<i>Score Fusion of CL,RS and EH</i>	127	0.71	2.66
<i>SF14</i>	<i>Score Fusion of CS,RS and EH</i>	120	2.86	4.52

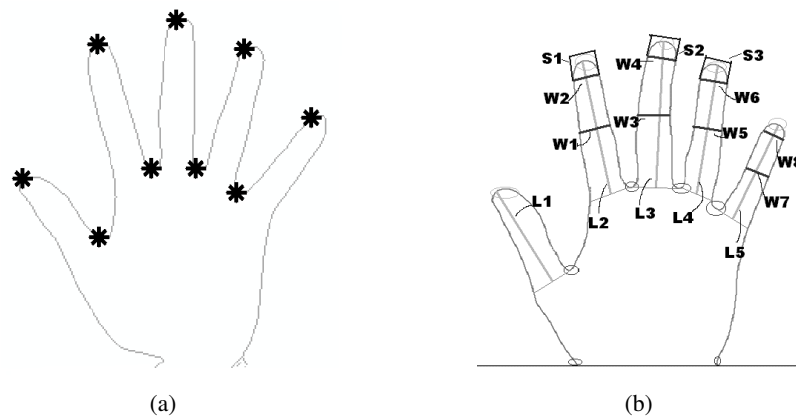


Figura 2.9: a) Pontos característicos da mão; b) Características utilizadas por Wong e Shi (2002).

O contorno é então analisado e, determinando os máximos e mínimos da sua localização espacial, é possível obter os pontos característicos da mão necessários à extração de características. Neste trabalho são utilizadas algumas das características utilizadas em Wong e Shi (2002) e são acrescentadas algumas novas. Todas elas serão descritas de seguida:

- Comprimento dos cinco dedos: mínimo, anelar, médio, indicador e polegar;
- Largura dos cinco dedos;
- Largura do pulso;
- Perímetro da mão;
- Área da mão;

Ao longo do processo de experimentação, o autor decidiu prescindir das seguintes características, face à sua baixa capacidade de discriminação:

- Comprimento do dedo polegar;
- Largura do dedo polegar;
- Largura do pulso;
- Área da mão;

Restando agora nove atributos para a classificação, esta foi feita recorrendo a três classificadores distintos. Um dos classificadores utilizados foi o KNN - *K-Nearest Neighbour Classifier*, que empregando duas medidas de distância obteve para o Erro Quadrático Médio (EQM) uma eficácia de 64.55% e para a Distância Média Absoluta (DMA) 73.64%. Outro classificador testado foi uma Rede Neuronal (RN) baseada em *Perceptron* multi-camada (Faundez-Zanuy et al., 2007). Este obteve a sua melhor *performance* usando três camadas utilizando a medida de distância do

EQM, com dez épocas, sendo esta de 93.64%. O último algoritmo testado neste sistema foi também uma RN baseada em *Radial Basis Function* (RBF) treinada de um modo discriminativo. A taxa de eficácia deste classificador foi de 90%.

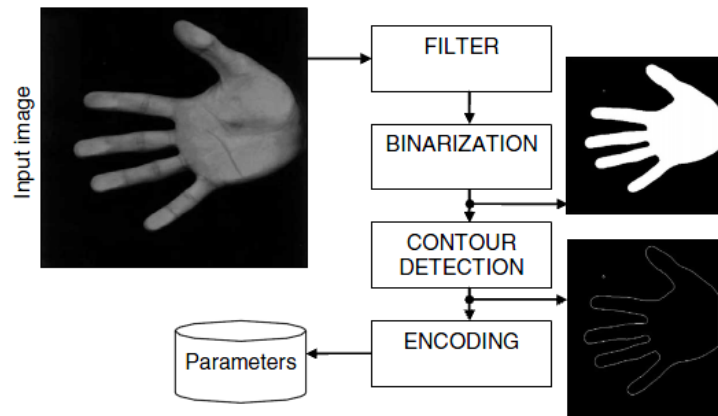


Figura 2.10: Algoritmo de pré-processamento usado por [Faundez-Zanuy e Merida \(2005\)](#).

Análise do contorno da mão recorrendo à transformação para coordenadas polares

Uma outra técnica passível de ser usada é a descrita no trabalho de [Adan et al. \(2008\)](#), que utiliza também um sistema de aquisição mais actual, ou seja, sem o recurso a fixadores, e fundamenta-se na representação em coordenadas polares do contorno da mão extraído após a segmentação da mesma. Neste trabalho são adquiridas a mão direita e esquerda da mesma pessoa, onde as características de ambas são concatenadas na fase de representação (treino) e também na fase de decisão (teste). Depois da extracção dos contornos, a sua transformação para coordenadas polares facilita a extracção dos pontos característicos da mão e, conseqüentemente, a extracção de características a serem usadas na identificação. As características utilizadas são uma série de medidas entre as pontas dos dedos e os vales entre estes. Foram atingidas neste trabalho uma taxa de eficácia de 96.5% e uma taxa de intrusão de 0.45%, recorrendo à técnica de medida de similaridade (das referidas distâncias) normalizada para um sistema de verificação. Para a identificação a taxa de eficácia alcançada foi de 97.5%.

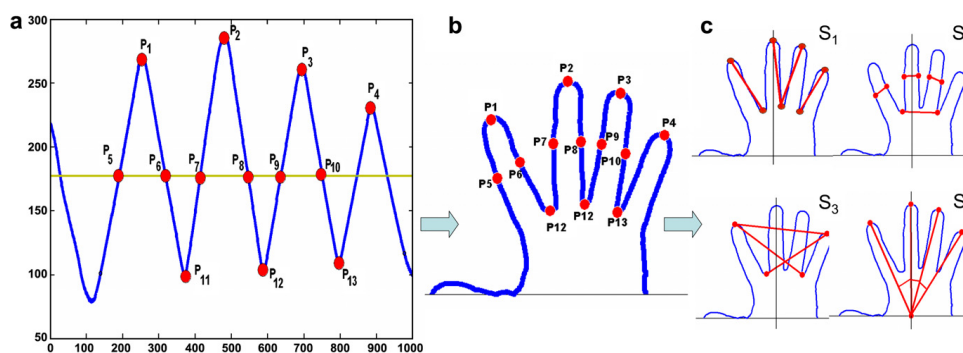


Figura 2.11: Método para detecção das características utilizado por [Adan et al. \(2008\)](#). a) Análise do contorno em coordenadas polares; b) Localização mais provável dos pontos detectados em (a) no contorno; c) Características utilizadas no reconhecimento.

Análise da curvatura do contorno da mão e segmentação dos dedos recorrendo a varrimentos na imagem

Mais recentemente, já em 2010, é apresentado um trabalho por [Montalvao et al. \(2010\)](#) com alguns avanços nas fases de pré-processamento e extracção de características. Este trabalho concentra-se, principalmente, na robustez no processamento de imagens quando adquiridas com baixa qualidade. O autor utiliza duas BD diferentes. Em ambas a aquisição foi feita sem fixadores, ainda que uma posição para a colocação da mão estava já pré-definida, não existindo assim uma liberdade total. O método de extracção dos pontos característicos neste trabalho baseia-se na análise da curvatura do contorno da mão, após a sua segmentação (ver Figura 2.12). No entanto, desde a segmentação até à extracção das características, existe sempre manipulação por parte do operador, o que constitui uma enorme desvantagem. Para a análise da curvatura, o operador tem de, primeiro, eliminar manualmente porções do pulso ou braço que possam eventualmente estar presentes na imagem do contorno e assim terminar o contorno ao nível do pulso. De seguida, é também necessário estimar que parte do contorno pertence a cada dedo para que cada um deles possa ser rodado e transladado devidamente para a extracção das medidas correspondentes.

Posteriormente, neste trabalho foi proposto um método dividido em quatro fases (Figura 2.13). A primeira trata de efectuar varrimentos horizontais na imagem e determinar os níveis de transição da escala de cinzento que correspondem aos dedos. Neste primeiro passo assume-se que se trata de uma imagem da mão esquerda disposta na vertical e que possíveis rotações da mão irão, certamente, influenciar as medidas extraídas. Num segundo passo é estimada a média da largura dos dedos e a delimitação da região da mão, sendo depois realizada a segmentação da palma e dos cinco dedos. Numa terceira fase deste método são estimados os comprimentos e larguras dos 5 dedos. Finalmente é feita uma comparação das medidas obtidas recorrendo à diferença de distâncias

de *Manhattan*. Utilizando este método, o autor conseguiu alcançar um *Equal Error Rate* (EER) de 3%.

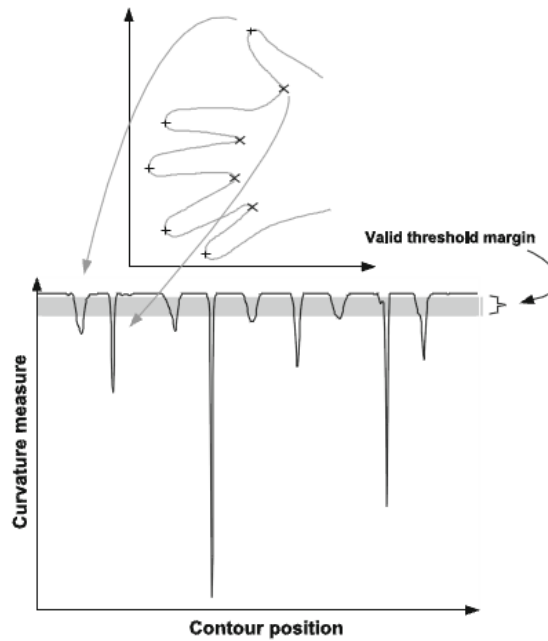


Figura 2.12: Método para detecção dos pontos característicos baseado na análise da curvatura do contorno (Montalvao et al., 2010).

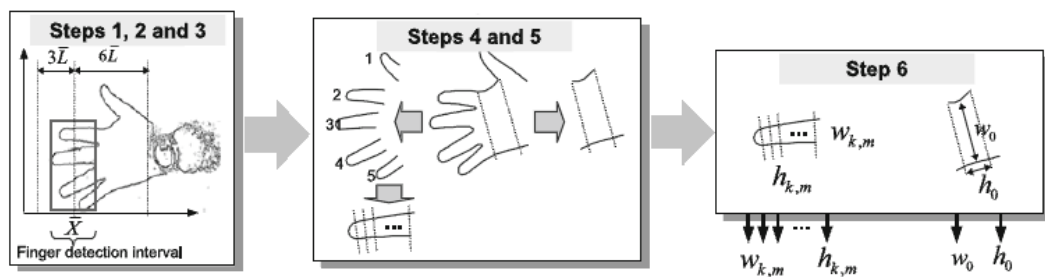


Figura 2.13: Método proposto por Montalvao et al. (2010) e suas fases constituintes.

2.3 Limitações dos Sistemas Baseados na Geometria da Mão

Os sistemas de reconhecimento baseados na geometria da mão apresentam as suas maiores limitações ao nível da aquisição e do desempenho.

Ao nível da aquisição existe o problema do conforto aquando da mesma, pois um sistema em que existam pinos a obrigarem o utilizador a colocar a mão de uma determinada forma não se mostra prático e os sistemas que vão surgindo sem tais fixadores apresentam ainda inúmeras restrições que quando não são respeitadas resultam na falha da autenticação. Isto traduz-se também numa diminuição do aspecto prático do sistema, devido ao facto de o utilizador poder ter de colocar a mão várias vezes até ser aceite.

Ao nível da performance e BD existe o obstáculo dos algoritmos pesados e lentos. Isto deve-se muitas vezes também a uma má estruturação da BD do sistema. Um dos aspectos mais relevantes a melhorar em termos de algoritmia é a robustez contra diversos factores resultantes de uma fraca aquisição, seja por um mau posicionamento da mão, por um maior relaxamento da mesma, por uma deficiência, por más condições de iluminação, por transpiração da mão, etc. O sistema deve, por isso, ser o mais imune possível a este tipo de condições.

Capítulo 3

Aquisição de Imagens para o Sistema de Reconhecimento

3.1 Introdução

Tal como pode ser constatado no capítulo anterior, existem grandes limitações na fase de aquisição deste tipo de sistemas. São impostas algumas restrições ao nível do posicionamento e configuração da mão e, dado que todos os sistemas trabalham dentro deste conjunto de limitações, as BD existentes estão todas sujeitas a tais restrições. A não existência de uma BD livre de restrições não potencia o desenvolvimento de investigação nesta área. Neste trabalho pretende-se eliminar grande parte, senão a totalidade, das restrições na aquisição. Foi então primordial a criação de uma nova BD de imagens da mão. A ideia principal passou pela criação de uma BD sem qualquer tipo de restrição na aquisição, de modo a fortalecer e a motivar o estudo, a investigação e o desenvolvimento de sistemas mais simples, mais práticos, mais cómodos e intuitivos.

3.2 Tipo de Aquisição

Pretende-se, nesta primeira fase do sistema, criar uma BD isenta de restrições. Foi então decidido fazê-lo o mais diversa possível e com uma ampla gama de aplicações possíveis. Tal como é indicado no trabalho de [Wong e Shi \(2002\)](#), um sistema de reconhecimento baseado na mão pode combinar as duas principais abordagens desta matéria, ou seja, basear o reconhecimento não só na geometria, mas também na impressão palmar. No entanto, para tal ser possível é necessário garantir uma resolução mínima nas imagens adquiridas para que seja exequível a extracção de características da palma da mão (linhas principais, textura, rugosidades, cristas, etc.). Foi assim escolhido e adquirido, para a aquisição, um scanner comercial, o EPSON Perfection V300 Photo, Figura 3.1.



Figura 3.1: Scanner utilizado na fase de aquisição.

As imagens foram adquiridas com as seguintes características:

- Resolução: 200 ppp
- Formato: '.tiff'
- Cor: 24 bits

Foram adquiridas 10 imagens da mão direita de cada pessoa e, durante a aquisição, o indivíduo foi fortemente encorajado a variar a posição da mão, orientação e configuração dos dedos, não tendo sido utilizados nenhum tipo de fixadores ou marcas indicadoras. Desta forma, pode afirmar-se que se trata de uma BD bastante diversa, no que respeita à aquisição, e sem qualquer tipo de restrição. Podem até ser observadas aquisições bastante originais, o que é aqui considerado como um ponto forte, um incentivo no avanço desta matéria. Ainda que esta BD de imagens da mão seja, neste trabalho, utilizada apenas para propósitos da geometria, foi sempre garantido durante a aquisição que esta cumpre todos os requisitos para poder ser aplicada em sistemas de reconhecimento baseados na impressão palmar (*Palmprint*). Todos os detalhes e minuciosidades da palma da mão podem ser claramente identificados nestas imagens.

3.3 Base de Dados Criada

A BD criada contempla, até agora, 10 imagens de 46 indivíduos diferentes. Pretende-se, numa fase posterior do trabalho, uma ampliação da mesma aumentando o número de indivíduos. Foi sempre tido em conta, durante a criação da BD, que esta deveria ser o mais diversa e equilibrada possível a nível de utilizadores. Deste modo, as aquisições efectuadas correspondem a indivíduos com as seguintes características:

- Raça branca e negra;
- Idades compreendidas entre os 10 e os 75 anos;
- Sensivelmente metade das aquisições corresponde a pessoas do sexo masculino e a outra metade a pessoas do sexo feminino;
- Nacionalidades diversas: Portuguesa, Cabo-Verdiana, Argelina, Nigeriana, Chinesa, Indiana, Brasileira, Iraniana e Polaca.

Na Figura 3.2 é possível perceber o tipo de aquisição e a liberdade aqui existente. A pessoa não só pode colocar a mão livremente na plataforma de aquisição como também não tem nenhum impedimento de rodar a mão ou até de colocar os dedos na configuração que entender. Não existe também nenhum cuidado com o facto de existirem porções do braço na imagem adquirida.

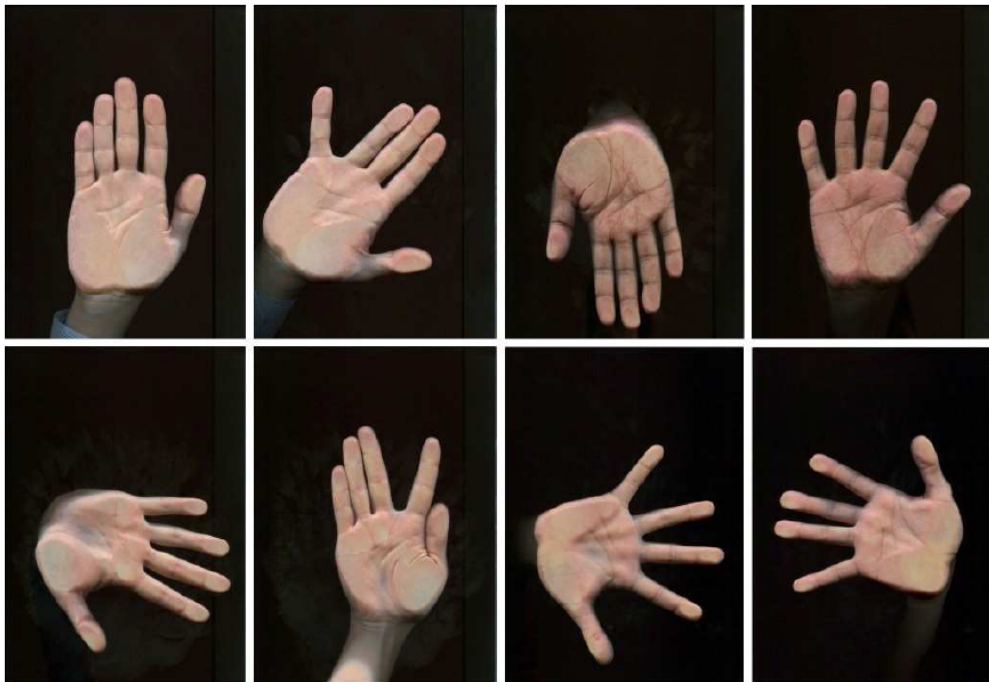


Figura 3.2: Alguns exemplos de imagens adquiridas (de notar a liberdade na posição, orientação e configuração da mão).

3.4 Conclusão

Foi apresentada uma nova base de dados que visa proporcionar o estudo e desenvolvimento de sistemas baseados na geometria da mão mais competitivos com os restantes tipos de reconhecimento. A única restrição que se colocou aos utilizadores foi que a aquisição deveria ser feita sem anéis. Os algoritmos desenvolvidos nas próximas fases terão este factor em consideração.

Desta forma, foi organizada uma competição internacional de algoritmos para detecção dos pontos característicos da mão, *Hand Geometric Points Detection*¹.

O desafio que esta competição coloca aos participantes consiste no desenvolvimento de um algoritmo otimizado para maximizar o rigor da detecção de nove pontos característicos da mão, capaz de mostrar um bom desempenho para universos maiores e também de eliminar as restrições impostas durante a aquisição. Ou seja, os algoritmos devem ser fiáveis independentemente do posicionamento/orientação da mão e da configuração dos dedos. Posteriormente, todos os algoritmos serão testados num conjunto de teste englobando todas estas características e vencerá o que possuir o maior índice de eficácia.

¹<http://paginas.fe.up.pt/hgc2011/>

Capítulo 4

Pré-processamento de Imagem

4.1 Introdução

Após a obtenção de uma imagem digital, a primeira etapa consiste na aplicação de algoritmos de pré-processamento. A função chave desta primeira etapa é melhorar e adequar a imagem às fases seguintes. Desta forma, e de acordo com o sistema aqui proposto, é necessário um tratamento prévio da imagem para posterior extração de características. Para tal é necessário saber quais os requisitos da fase seguinte que corresponde à detecção de pontos característicos. Sabemos então que a imagem original se apresenta no formato RGB (*Red, Green and Blue*) e que são necessárias a detecção e segmentação da mão nessa mesma imagem. Posteriormente, será feita uma binarização para obtenção da máscara binária da mão e consequente extração dos contornos da mesma. Na fase final do bloco de pré-processamento é feito o alinhamento da imagem dos contornos para de uma forma mais simples fazer a extração dos nove pontos característicos.

Imagens digitais, Matlab, e espaço de Cores RGB

Uma imagem digital consiste num conjunto de pixels dispostos numa matriz com um determinado tamanho em que cada elemento representa a cor de um pixel. A cor de cada pixel pode ser representada de diversas formas e isso depende do espaço de cores em que a imagem está mapeada. É comum que as fotografias digitais provenientes de meios como câmaras digitais, sistemas de aquisição de vídeo ou webcams estejam mapeadas no espaço de cor RGB. Neste espaço, a cor de cada pixel da imagem é determinada pelas componentes de vermelho, verde e azul que a compõem, pelo que o que a matriz contém em cada posição (x, y) são três valores que quantificam a quantidade de cada cor. A título de exemplo, e utilizando uma codificação de oito bits, uma cor vermelho puro teria um valor de 255 para a componente vermelho e zero para as componentes

verde e azul. Uma imagem digital é interpretada pelo software Matlab como uma matriz de dimensões $largura*altura*3$, em que a dimensão 3 corresponde às três componentes que definem a cor do pixel. Cada elemento contém o valor que a cor desse pixel toma para essa componente.

A Figura 4.1 ilustra o sistema de coordenadas de referência usado em processamento e análise de imagem (Gonzalez e Eddins, 2004).

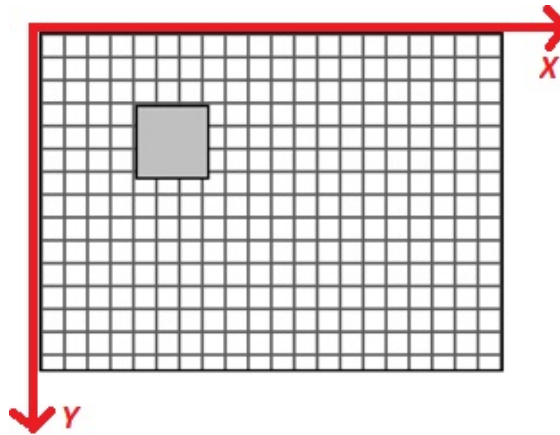


Figura 4.1: Sistema de coordenadas de referência usado.

4.2 Segmentação Binária

4.2.1 Introdução

As metodologias para a detecção de objectos na imagem são muitas vezes baseadas no histograma de propriedades, em geral, o nível de cinzento. Dado que o número de pixels numa imagem é normalmente muito elevado pode-se considerar o histograma como uma boa aproximação à densidade de probabilidade da propriedade que ele representa.

Binarização: *Thresholding* com um único limiar de separação. A selecção do limiar de separação (separação entre as classes "claro" e "escuro" no caso da propriedade ser o brilho) faz-se geralmente no vale do histograma, sendo este bimodal, tal como pode ser observado na Figura 4.2.

A localização do vale entre picos de um histograma bimodal não é, geralmente, muito fácil devido ao carácter discreto dos níveis do histograma e à presença de ruído nas imagens. Próximo do vale podem existir diversos mínimos locais, não sendo trivial encontrar automaticamente o limiar óptimo. Para resolver este problema é frequentemente útil proceder a uma suavização da imagem antes da determinação do histograma (por exemplo, usando filtros de média), ou a uma pós-filtragem do histograma (interpretado como um "sinal" unidimensional).

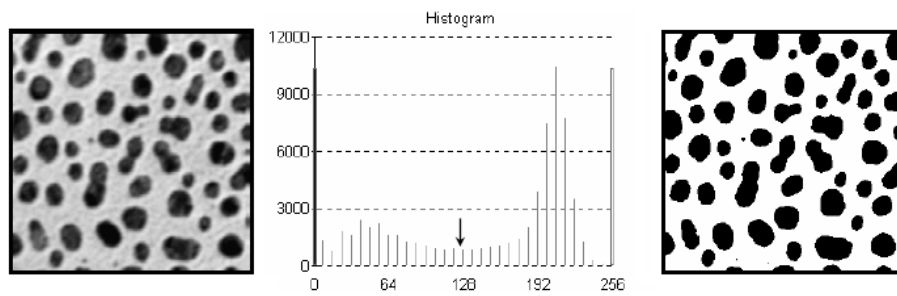


Figura 4.2: Exemplo de binarização a partir da definição de um valor de *threshold*.

Thresholding multi-nível - onde se consideram mais classes de pixels do que apenas claros e escuros (por exemplo quase brancos, cinzentos claros, cinzentos escuros, quase pretos). Idealmente para estes casos os respectivos histogramas devem ser multi-modais (quatro modas no exemplo anterior), procedendo-se à localização de múltiplos limiares de separação nos vales entre modas (três limiares no exemplo anterior).

Thresholding variável - uma outra forma de se proceder à binarização de uma imagem consiste em subdividi-la num mosaico de sub-imagens, determinando-se o histograma de cada uma dessas sub-imagens. Nas sub-imagens em que o histograma é bimodal determina-se o valor do limiar de separação apropriado. O conjunto de limiares obtido pode então ser usado para, por um método de interpolação, se determinar, para cada ponto da imagem original, o valor de limiar apropriado. O conjunto de todos os limiares pontuais constitui uma superfície de referência para binarização. Este método é muitas vezes útil, em especial quando a iluminação da cena é não é uniforme.

Tendo em vista a próxima etapa, a obtenção dos contornos da mão presente na imagem é essencial que exista uma clara demarcação da zona da mão e do fundo. Aqui será utilizado um *thresholding* com apenas um limiar de separação - binarização, pois apenas se pretende a segmentação de um único objecto. O resultado esperado será obter uma clara definição da mão, em que os pixels correspondentes ao fundo sejam representados pelo nível correspondente ao branco e os restantes pelo nível correspondente ao preto. Após esta operação serão aplicados métodos de eliminação de ruído para a obtenção de uma imagem final onde a máscara binária da mão seja óbvia e bem definida.

4.2.2 Métodos e Resultados

Método I - Valores de *Hue* da imagem

Tendo por base o conhecimento *à priori* do que se pretende detectar, pode ser usado um método de binarização baseado nos tons de matiz da pele na imagem. Este método pratica uma detecção orientada ao pixel e baseia-se no facto de o tom de pele de todos os humanos ter um matiz muito semelhante. Cedo se compreendeu que o espaço de cor RGB não é o mais indicado para este objectivo, dado que a cor da pele toma valores espalhados pela gama completa de cada um dos três canais. Optou-se então pelo espaço de cor HSV (*Hue*, *Saturation*, *Value*), no qual a cor de cada pixel é dada pelos valores de matiz (*Hue*), saturação (*Saturation*), e brilho (*Value*).

Numa primeira fase foi efectuada uma recolha e selecção de amostras de fotografias com porções da pele, ao que se seguiu o respectivo cálculo do valor de matiz. Após obtenção do intervalo de níveis de matiz que os tons referentes à pele tomam, partiu-se para o desenvolvimento do sistema em si. A imagem original, lida em RGB, é inicialmente mapeada para o espaço de cor HSV. De acordo com os níveis do matiz (*Hue*) calculado na fase anterior, foi criada uma máscara binária com os pixels referentes à pele a preto e todos os restantes a branco. A Figura 4.3 ilustra o efeito de cada parâmetro do espaço de cor HSV na cor de um dado pixel. O valor de matiz (*Hue*) corresponde à cor do pixel, o valor da saturação (*saturation*) à pureza dessa cor, e o valor do brilho (*value*) à intensidade - energia - dessa cor.

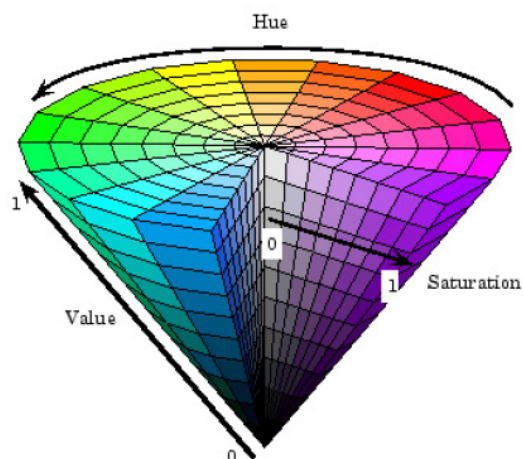


Figura 4.3: Espaço de cores HSV (<http://clubelek.insa-lyon.fr/ressources/video/index2.php>).

Dado que o matiz da cor da pele assumiu sempre um valor situado num curto intervalo, o mapeamento de uma fotografia para o espaço de cor HSV permitiu o reconhecimento do tom de pele através da análise de um único parâmetro (ao invés de estar a analisar as três componentes de cor do espaço RGB) e ainda uma certa independência das condições de iluminação em que a imagem foi capturada (dado que essas vão influenciar, idealmente, apenas o brilho).

A primeira operação foi então convertê-la do espaço de cor RGB para HSV, tirando partido da função da *toolbox* de processamento de imagem do Matlab já existente para o efeito: *rgb2hsv()*. De seguida são separadas as três componentes e copiado o matiz para uma matriz isolada para análise posterior. Na Figura 4.4(a) está representada uma imagem original no espaço de cores RGB e na Figura 4.4(b) a representação dos seus valores do matiz (canal *Hue*) em escala de cinzas.

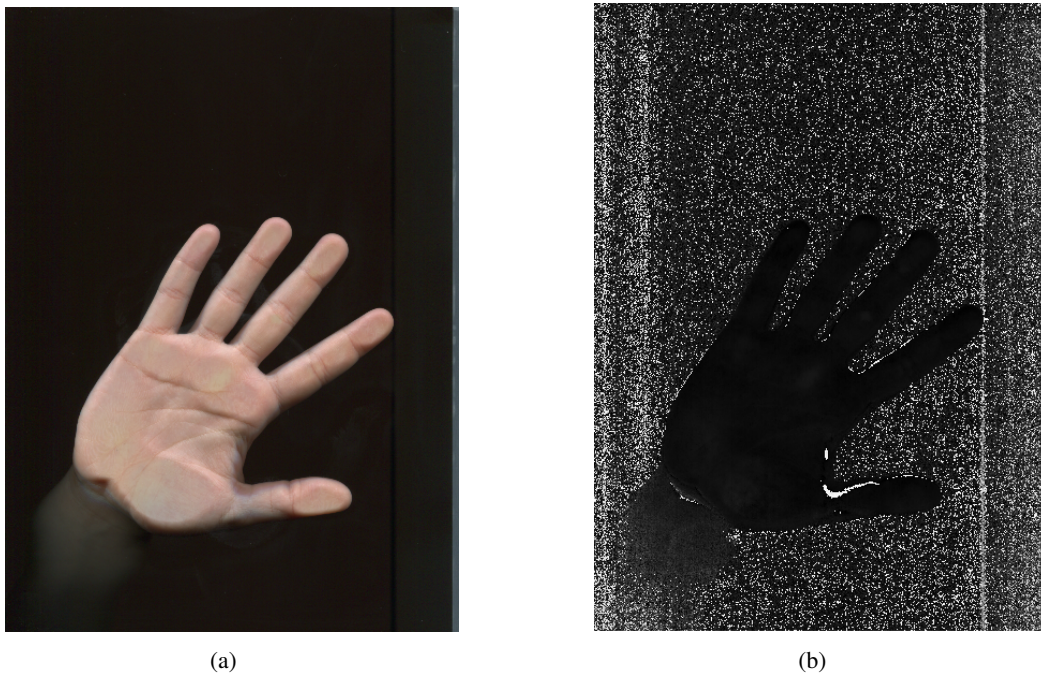


Figura 4.4: a) Imagem original no espaço de cores RGB, b) Representação do canal *Hue*, em escala de cinzento, da imagem a).

Identificação do intervalo de níveis do matiz correspondente à cor da pele

Foi então estabelecido que se iriam utilizar unicamente os valores do matiz para identificar pixels correspondentes à cor da pele. Cada uma das imagens anteriormente seleccionadas foi processada em Matlab: foi calculada a matriz representante do seu matiz (sendo o valor zero correspondente ao valor de matiz mínimo e o valor um correspondente ao valor máximo) e obtido o histograma desse componente (que, para cada valor do matiz, indica a quantidade de pixels na imagem com esse valor). A título exemplificativo, são apresentadas na Figura 4.5 duas dessas amostras e respectivos histogramas do matiz:

Através da repetição deste processo para um determinado número de amostras (5 amostras, cada uma contendo pele de várias zonas diferentes) revelou-se notório que o matiz da cor da pele ocupa uma faixa estreita nos 10% inferiores do alcance completo do matiz (de 0 a 1, na

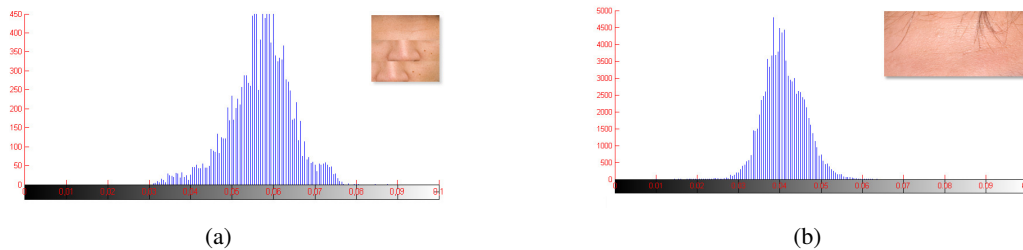


Figura 4.5: a) e b) Histogramas representativos da distribuição dos valores do matiz em duas amostras distintas.

representação utilizada). Em detalhe, aparenta ser uma distribuição aproximadamente gaussiana centrada entre 0,04 e 0,06. Este valor foi posteriormente refinado através de testes com diferentes imagens, e chegou-se a uma gama de [0.0055, 0.0755]. Foi assumido então que todos os pixels com um matiz superior a 0.0055 e inferior a 0.0755 representam pele. Se estes valores funcionam bem para a maioria dos casos, poderão existir outros em que um pequeno ajuste é necessário.

Criação da máscara binária de tons da pele

Realizado o estudo prévio que permitiu conhecer os valores que o matiz da cor da pele assume, iniciou-se a estruturação do algoritmo propriamente dito. Foi implementado um algoritmo em Matlab que recebia como argumentos a imagem original e os níveis de matiz, entre os quais se considera que um determinado pixel pertence à pele. O algoritmo percorre todos os pixels da imagem original e para cada um deles escreve zero ou um no pixel correspondente de uma outra imagem, conforme o matiz da imagem original esteja dentro ou fora do *threshold* considerado para a cor da pele. No final do algoritmo obtém-se uma nova imagem binária onde cada pixel tem uma correspondência na imagem original. Nesta imagem, pixels detectados como pertencentes à pele aparecem a preto. A Figura 4.6 é ilustrativa de um dos primeiros resultados obtidos com a função (correndo apenas o algoritmo descrito e com o *threshold* fixo).

O local referente à mão encontra-se claramente demarcado, contudo a imagem resultante encontra-se fortemente preenchida com ruído. Para pixels isolados e pequenos aglomerados de pixels que aparecem na máscara inicialmente, foi aplicado um algoritmo de etiquetagem de componentes na imagem e foram eliminados todos os aglomerados de pixels cuja área fosse inferior a um determinado valor. No entanto, acabou por se utilizar uma aproximação mais simples e computacionalmente mais leve: aplicou-se à máscara binária um filtro de ordem (mediana), fazendo com que pixels isolados e pequenos aglomerados de pixels se "fundissem" com o fundo preto e assumissem uma tonalidade cinzenta (média de pixels brancos com preto circundante), mas que grandes aglomerados (como o caso da mão) mantivessem a cor branca. De seguida aplicou-se a essa imagem um *threshold* em escala de cinzentos em que todos os valores de cinzento abaixo de

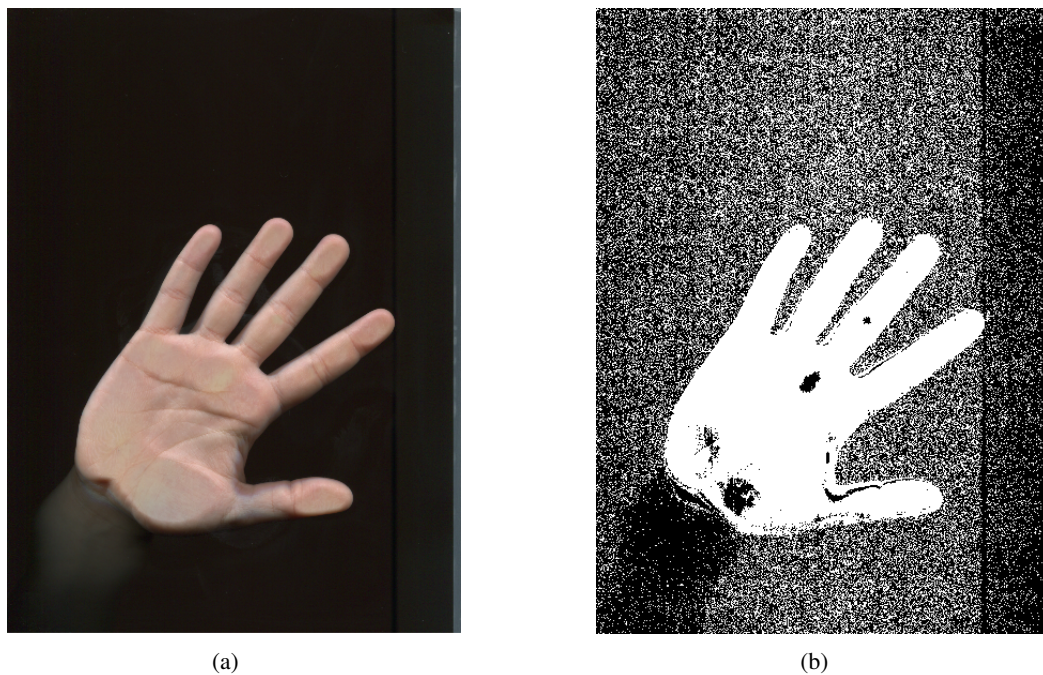


Figura 4.6: a) Imagem original no espaço de cores RGB, b) Primeiro ensaio da obtenção da máscara binária correspondente à imagem em a).

um determinado nível (0.7, neste caso) se tornassem pretos, obtendo-se uma imagem muito mais limpa de ruído. Através deste método e da afinação do *threshold* para a cor da pele, conseguiu-se a criação de máscaras binárias de qualidade muito aceitável, como se ilustra na Figura 4.7:

É notória a melhoria da qualidade da máscara binária através da aplicação da estratégia descrita anteriormente: a maioria dos pontos isolados e do ruído desapareceu, mas a zona correspondente à mão continua bem identificada. O exemplo acima mostra que o sistema é capaz de funcionar em condições variadas, com um certo grau de independência da iluminação dos fundos utilizados e da raça da pessoa (os matizes da cor da pele das várias raças são muito semelhantes). Como pode ser verificado pela imagem da máscara binária obtida, mesmo após aplicação dos filtros, permanecem ainda alguns "buracos" no interior da mão. Este problema poderia ser resolvido ajustando os valores dos filtros aplicados. No entanto, tal alteração iria prejudicar a binarização de uma forma global. Isto é, para fechar os buracos no interior da mão, o espaço entre os dedos iria ser também progressivamente preenchido. Deste modo, recorreu-se a um algoritmo de etiquetagem de componentes para resolver este problema.

Para efeitos de análise, é importante discriminar entre os diversos objectos presentes na máscara binária, atribuindo a todos os pixels de um objecto uma etiqueta exclusiva. Para se conseguir essa etiquetagem (*component labeling*) pode ser usado o seguinte algoritmo básico:

1. Efectuar um varrimento da imagem;
2. Para cada pixel de valor um encontrado (X), analisar os vizinhos imediatos já visitados;

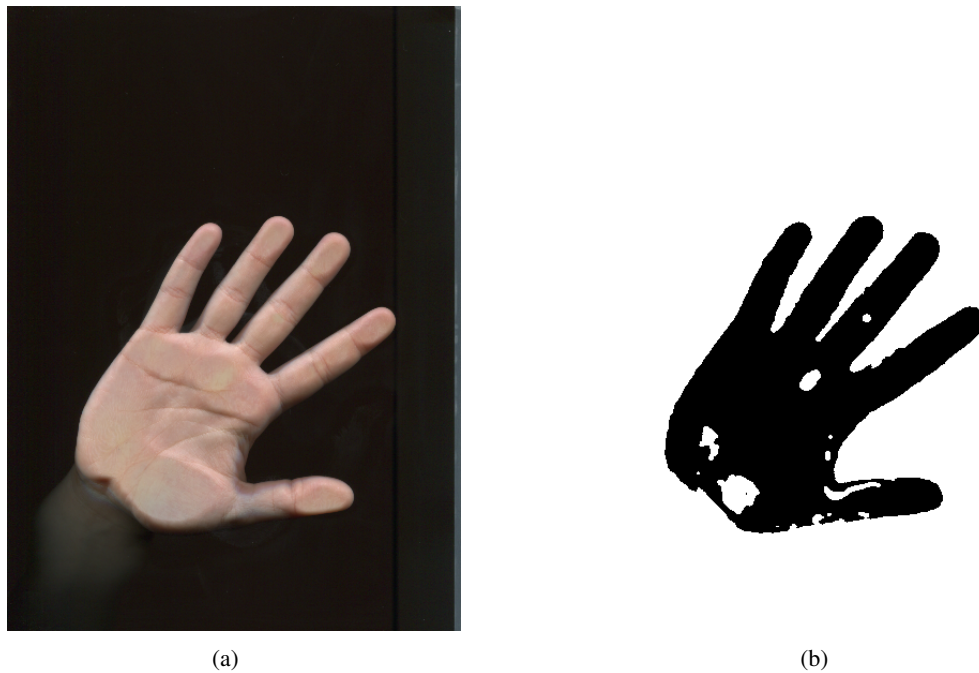


Figura 4.7: a) Imagem original no espaço de cores RGB, b) Segundo ensaio da obtenção da máscara binária correspondente à imagem em a), agora com aplicação de alguns filtros.

3. Se todos esses vizinhos forem zero, atribuir a X uma etiqueta nova;
4. Caso contrário, lendo as etiquetas dos vizinhos por determinada ordem, atribuir a X a primeira etiqueta diferente de zero;
5. Se, entre os pontos vizinhos houver mais do que uma etiqueta, assinalar a equivalência dessas etiquetas numa lista;
6. Terminado o varrimento, processar a lista de equivalências, definindo uma etiqueta única para cada grupo de etiquetas equivalentes;
7. Efectuar um segundo varrimento da imagem, atribuindo aos pixels as novas etiquetas.

O Matlab contém a função descrita acima na sua *toolbox* de processamento de imagem, *bwconncomp()*. Esta função retorna o número de objectos diferentes encontrados, a lista dos respectivos índices na imagem e a área de cada objecto. Recorrendo a este último parâmetro, a área de cada objecto encontrado, é possível identificar e seleccionar a mão eliminando os restantes objectos (ver Figura 4.8).

Os resultados obtidos parecem bastante satisfatórios. No entanto, quando este método foi testado em todas as imagens da BD verificou-se que para determinadas imagens não foi possível obter o resultado esperado, nomeadamente para aquelas imagens em que dois ou mais dedos estão juntos e também para imagens em que a mão apresenta um aspecto mais rosado. Para o segundo caso, uma variação dos níveis de *threshold* e dos filtros foi experimentada mas em consequência

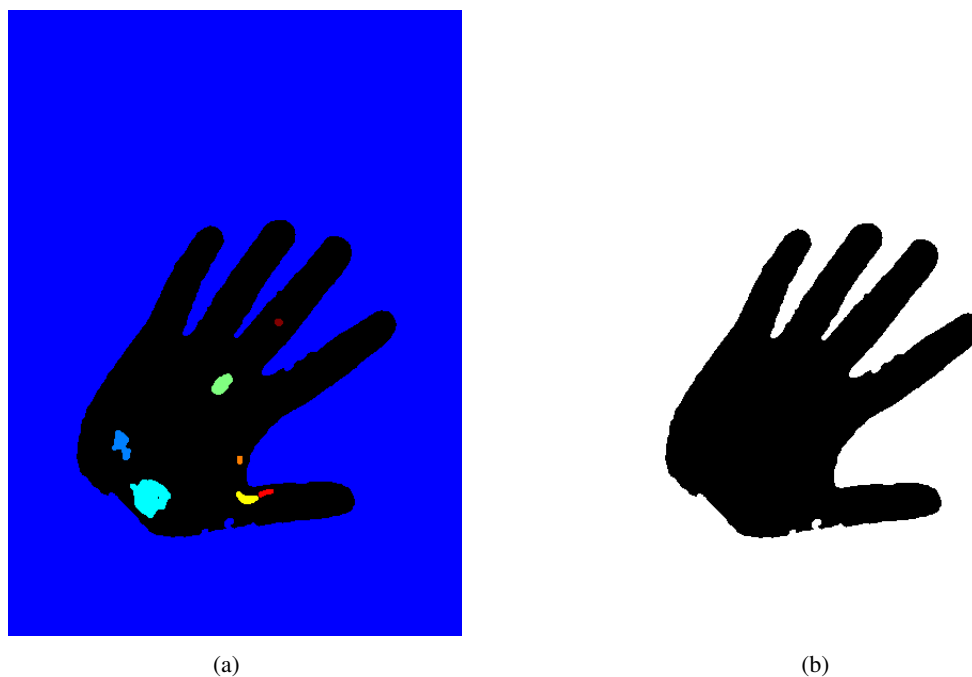


Figura 4.8: a) Atribuição de uma etiqueta diferente a cada objecto separado dos restantes; b) Detecção do objecto com maior área e eliminação dos restantes - Máscara binária final.

da melhoria desse aspecto, outros eram afectados negativamente, como por exemplo, os vales ganhavam formas bastante irregulares (cristas) que influenciariam bastante a detecção dos pontos correspondentes aos vales. Tais resultados podem ser observados na Figura 4.9 e Figura 4.10:

Método II - Função *Ridge-Segments()* (Kovesi)

Devido às limitações apresentadas pela função de binarização anteriormente apresentada, e no sentido de combater os problemas encontrados foi feita uma pesquisa de outros métodos de binarização. Durante essa pesquisa surgiu o interesse em testar métodos de realce do contraste para verificar se, de alguma forma, tal ajudaria a uma melhor definição do *threshold* de binarização.

Normalização

O principal objectivo deste passo é o reforço do contraste na imagem original, de modo a reduzir variações nos níveis de cinzento da imagem e assim facilitar a segmentação da área pretendida de forma eficaz (Magalhaes et al., 2009).

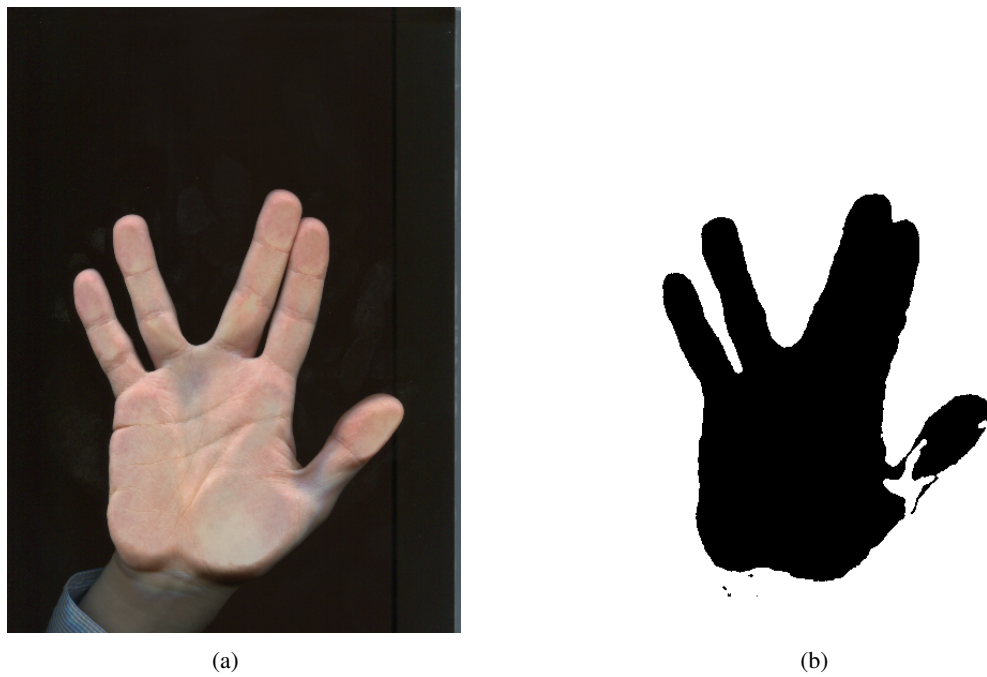


Figura 4.9: a) Imagem original com dedos juntos; b) Respectiva máscara binária com falha na separação de todos os dedos.

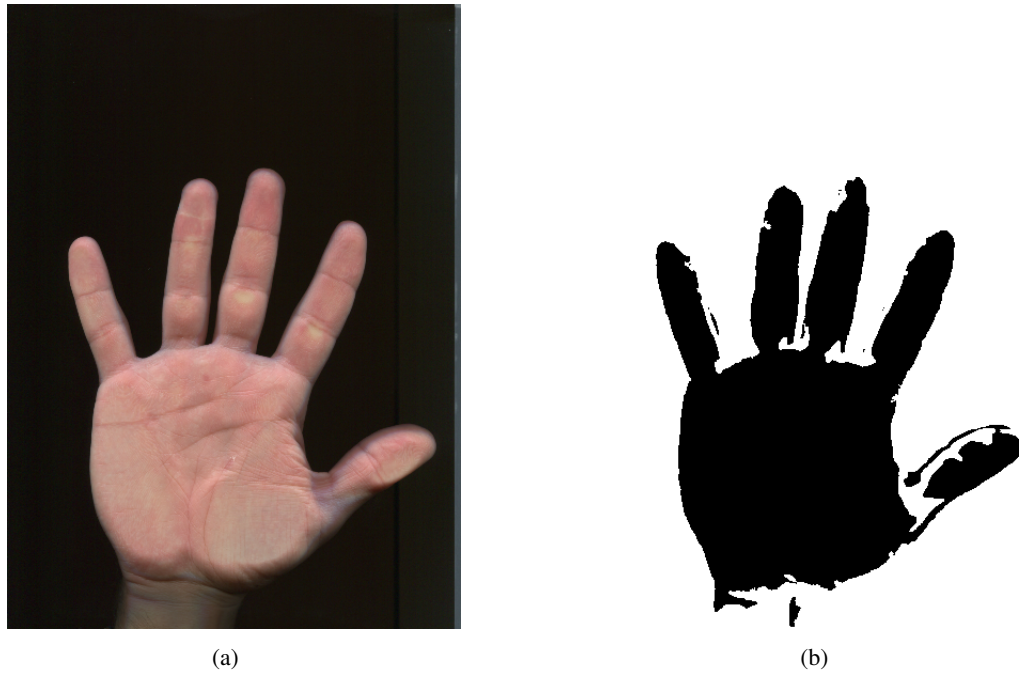


Figura 4.10: a) Imagem original com tonalidade mais rosada; b) Respectiva máscara binária mostrando deficiências nas extremidades da mão e as "cristas" nos vales.

Seja I uma imagem em tons de cinzento e $I(x, y)$ o valor do pixel correspondente às coordenadas (x, y) , μ representa a média estimada e σ o desvio padrão estimado. A normalização é então aplicada à imagem usando a seguinte equação:

$$I_n(x, y) = \begin{cases} \mu_n + \sqrt{\frac{\sigma_n(I(x, y) - \mu)^2}{\sigma}}, & \text{se } I(x, y) > \mu \\ \mu_n - \sqrt{\frac{\sigma_n(I(x, y) - \mu)^2}{\sigma}}, & \text{em outro caso} \end{cases} \quad (4.1)$$

onde $I_n(x, y)$ representa o valor de cada pixel em escala de cinzentos normalizado. μ_n e σ_n representam assim a média e desvio padrão, desejados.

Descrição da Função Utilizada

De forma a implementar o método acima descrito foi aplicada a função *Ridge-Segments()* utilizada por Kovesi¹, anteriormente aplicada em imagens de impressões digitais para evidenciação de cristas. A função, *RidgeSegments()*, normaliza os valores de intensidade da imagem para realçar regiões de elevado contraste (idealmente, a zona da mão) com a finalidade de que a região de interesse tenha μ (média) zero e σ (desvio padrão) unitário. Para tal, a função efectua o seguinte procedimento:

1. Parte a imagem original em vários blocos ($N \times N$) (N definido à chamada da função)
2. Avalia o desvio padrão em cada um dos blocos:
 - SE: desvio padrão maior do que o nível de *threshold*
 - ENTÃO: bloco pertence à região de interesse
 - SE: desvio padrão menor do que o nível de *threshold*
 - ENTÃO: bloco não pertence à região de interesse

Parâmetros de entrada da função descrita:

- Imagem a ser segmentada;
- Dimensão do bloco sobre o qual é determinado o desvio padrão;
- *Threshold* do desvio padrão.

Retorna:

- Imagem normalizada, com μ 0 e σ unitário;

¹<http://www.csse.uwa.edu.au/~pk/>

- Máscara binária indicadora de regiões de cristas, valendo um na região de interesse e zero no restante;

Resultados

Após implementação da função acima descrita seguiu-se a fase de testes da mesma com as imagens da BD. Começando por se fazer um ajuste dos parâmetros de entrada da função (nível de *threshold* e tamanho dos blocos). Tal como pode ser verificado na Figura 4.11 e Figura 4.12, uma correcta e bem definida segmentação da zona da mão foi conseguida. Tanto em imagens com os dedos afastados como em imagens que apresentam dois ou mais dedos juntos, este método mostra-se capaz de separar os dedos.

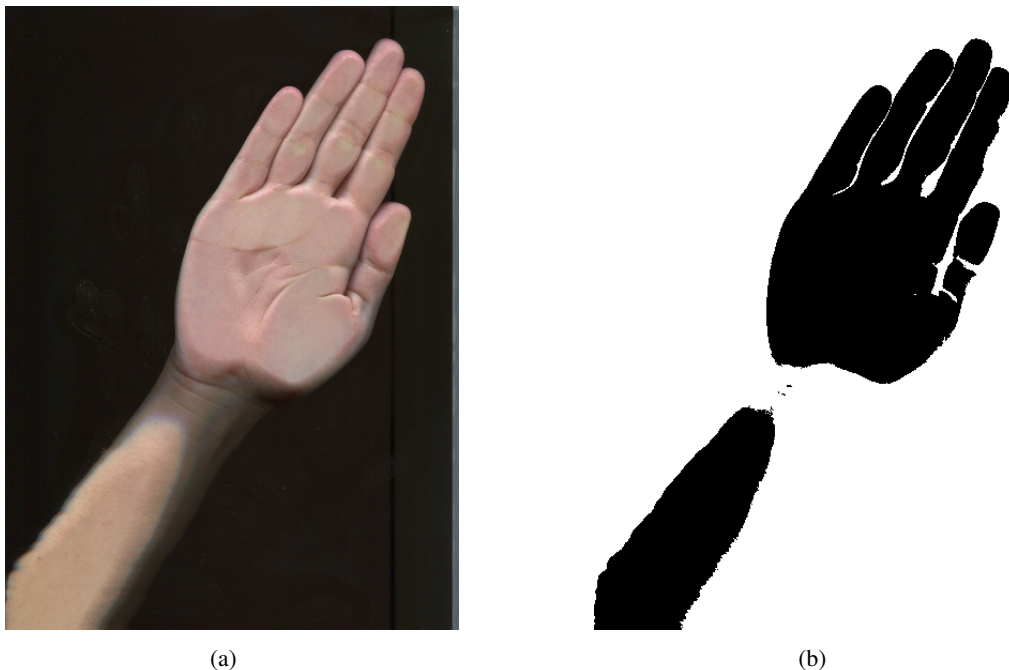


Figura 4.11: a) Imagem original com todos os dedos completamente juntos; b) Respectiva máscara binária definindo claramente a zona entre os dedos.

No entanto, alguns resultados negativos foram também identificados. Podemos observar casos em que a alta sensibilidade deste método, que permite a separação dos dedos quando estes se encontram juntos, faz também com que pequenas sombras sejam suficientes para separar duas zonas da mão, nomeadamente, provocar a desconexão dos dedos (ver Figura 4.13 e Figura 4.14). Note-se que não se trata de uma má aquisição, apenas de pequenos sombreados que estão muitas vezes presentes devido a propriedades anatómicas do utilizador. Registou-se também que para indivíduos de raça negra este problema agrava-se devido à forte evidência das dobras dos dedos.

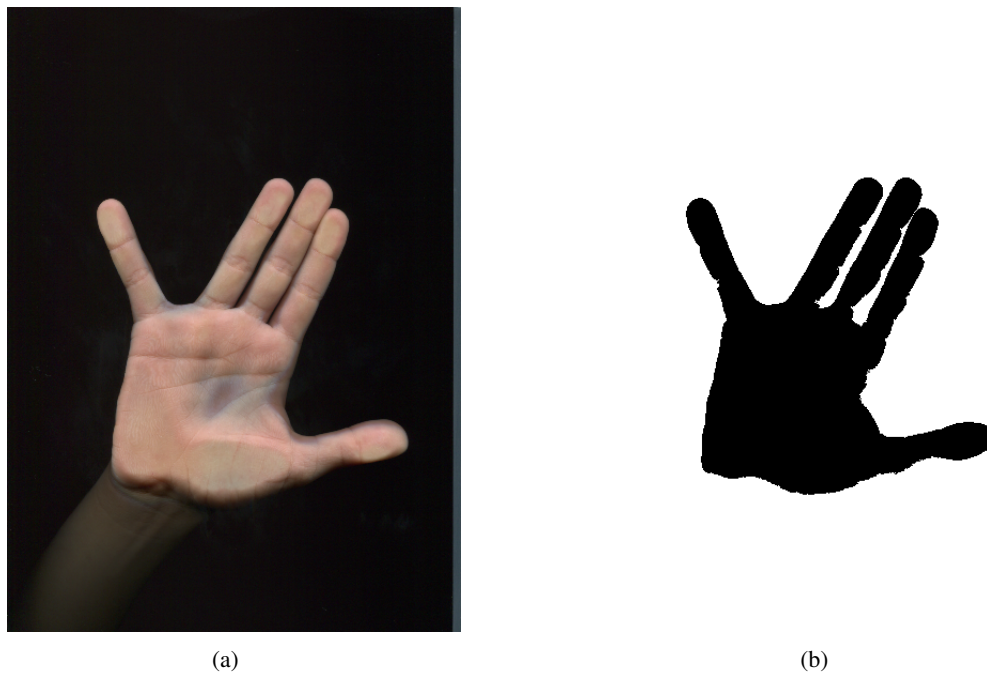


Figura 4.12: a) Imagem original com alguns dedos juntos; b) Respectiva máscara binária definindo claramente a zona entre os dedos.

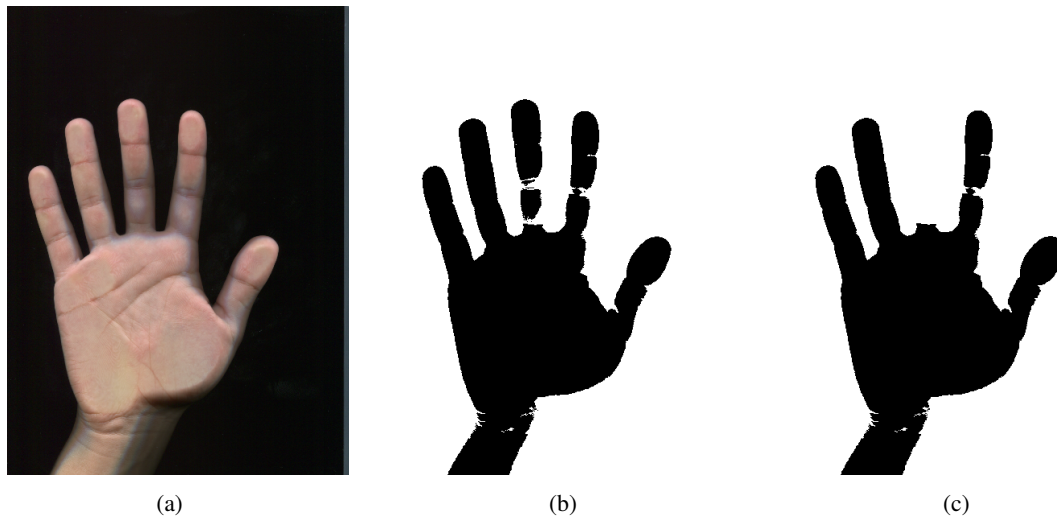


Figura 4.13: a) Imagem original com todos os dedos abertos e uma ligeira sombra nos dedos indicador e médio; b) Respectiva máscara binária (de notar que existe mais que um objecto); c) Máscara binária após etiquetagem dos vários componentes.

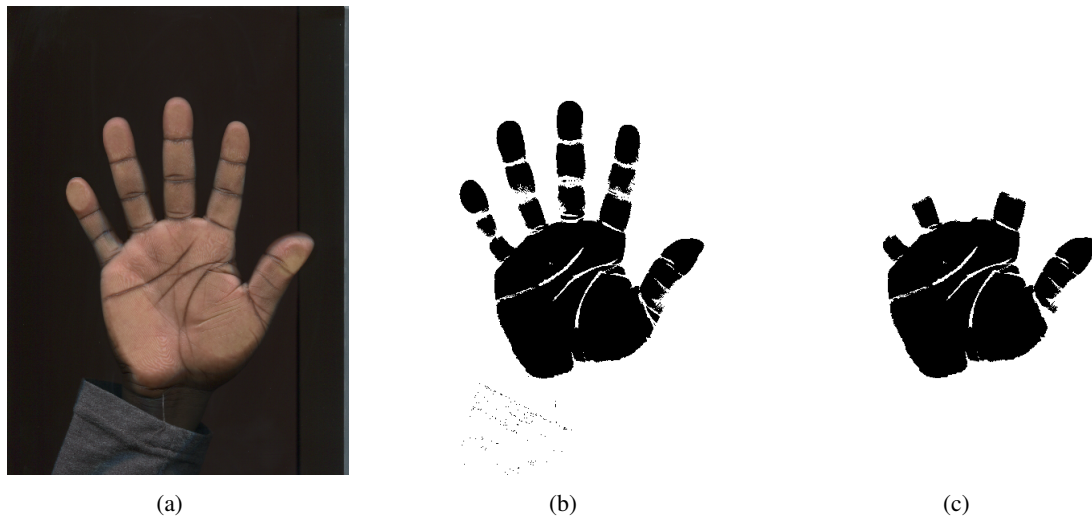


Figura 4.14: a) Imagem original de indivíduo de raça negra com todos os dedos afastados; b) Respectiva máscara binária (de notar que existe mais que um objecto); c) Máscara binária após etiquetagem dos vários componentes.

Método III - *Thresholding* Dinâmico - Método de Otsu

O método de Otsu é uma técnica de *thresholding* global para escolher o melhor *threshold* a usar para a binarização da imagem, sendo muitas vezes chamado de *thresholding* dinâmico (Liu e Yu, 2009). Este método baseia-se no histograma normalizado (da imagem) como uma função de densidade de probabilidade discreta, da seguinte maneira:

$$P_r(r_q) = (n_q/n), n = 1, 2, 3, \dots, L-1 \quad (4.2)$$

onde n é o número total de pixels na imagem, n_q é o número de pixels que tem intensidade r_q e L é o número total de possíveis níveis de intensidade na imagem. O método de Otsu escolhe o *threshold* de valor k (tal que k é um nível de intensidade onde $C_0 = [0, 1, \dots, k-1]$ e $C_1 = [k, k+1, \dots, L-1]$) que maximiza a variância σ^2 entre classes, que é definida como:

$$\sigma^2 = P_0(\mu_0 - \mu_T)^2 + P_1(\mu_1 - \mu_T)^2 \quad (4.3)$$

onde, $P_{0,1}$ corresponde às probabilidades das classes e $\mu_{0,1}$ à respectiva média.

Em imagens digitais, a uniformidade dos objectos tem um papel importante na separação destes objectos do fundo. De facto, a abordagem de *Otsu* para a binarização de imagens em tons de cinzento é eficiente com base na uniformidade medida entre as duas classes C_0 e C_1 a serem segmentadas. O método proposto caracteriza-se pela sua natureza não paramétrica e não supervisionada de selecção do *threshold* e possui as seguintes vantagens:

- O processo como um todo é muito simples; são utilizados somente os momentos acumulativos zero e de primeira ordem do histograma de níveis de cinzento;
- Uma extensão directa para problemas que exijam vários limiares é viabilizada face ao critério no qual o método está baseado, seja qual for a análise de discriminantes;
- Um limiar óptimo (ou conjunto de limiares) é seleccionado de forma automática e estável, não baseado na diferenciação (uma propriedade local como um vale) mas sim na integração (propriedade global) do histograma;
- Viabiliza a análise de outros aspectos importantes, tais como a estimativa dos níveis médios das classes, avaliação da separabilidade das classes, etc;
- O método é extremamente genérico; cobre um vasto leque de processos de decisão não supervisionados; o conjunto das suas aplicações não se limita à binarização de imagens em níveis de cinza tal como descrito anteriormente, mas pode também ser utilizado noutros casos de classificação não supervisionada nos quais um histograma de alguma característica discriminativa que classifique objectos esteja disponibilizado.

Cabe ressaltar no entanto, que o limiar calculado não é apropriado para imagens em que as variações dos níveis de cinzento de objectos e fundo ou populações de pixels correspondentes sejam extremamente diferentes.

Resultados

A *toolbox* de processamento de imagem presente no MatLab contém já uma função que implementa este método de binarização, a função *graythresh()*. Este método, além de fácil e rápido de implementar, apresentou um bom desempenho na binarização das imagens, revelando-se bastante leve computacionalmente. Existe também na *toolbox* de processamento de imagem uma função, *imadjust()*, que ajusta a intensidade da imagem, resultando num aumento do contraste da imagem original em tons de cinzento, (Figura 4.15). Esta função foi utilizada no sentido de proporcionar melhores resultados à função *graythresh()*.

Este método apresentou bons resultados, em geral, melhores do que os dos métodos anteriores, (Figura 4.16). Nas imagens onde se observam pequenas sombras, este método consegue separar correctamente a zona da mão sem que alguma parte dela (por exemplo, o dedo) seja deixada de parte. Tal pode ser verificado na Figura 4.17. Contudo, em situações em que a mão se encontra com os dedos juntos este método apresenta algumas dificuldades para a correcta binarização da imagem. Embora seja capaz de segmentar correctamente, em alguns desses casos (Figura 4.18), noutros tal não acontece (Figura 4.19). Um último teste deste método foi feito recorrendo a imagens de indivíduos de raça negra e apesar de apresentar melhores resultados do que o método anteriormente descrito (Método II), continua ainda assim a mostrar falhas em algumas das imagens (Figura 4.20).

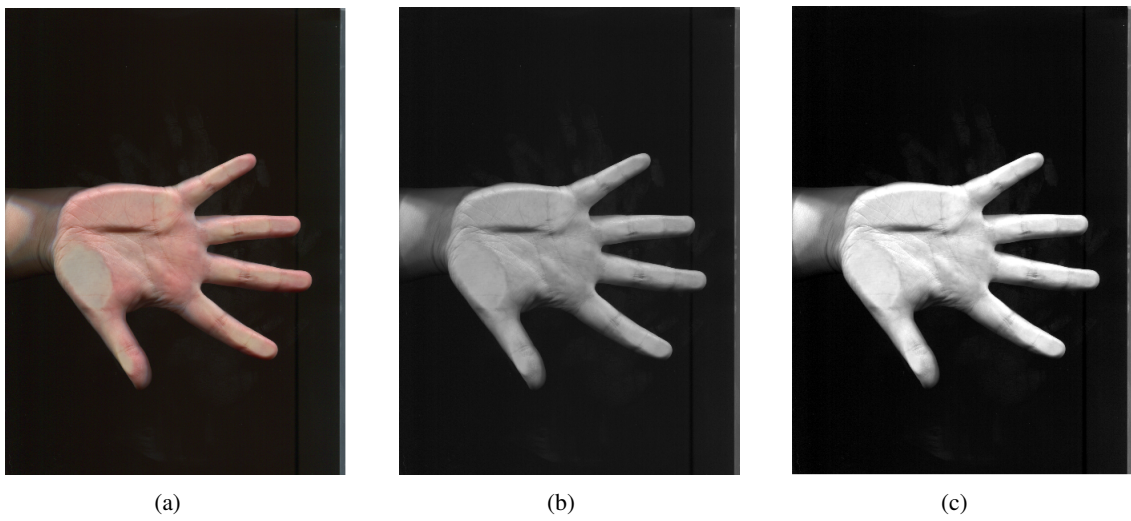


Figura 4.15: a) Imagem original RGB; b) Imagem original em tons de cinzento; c) Imagem original em tons de cinzento, resultado da aplicação de *imadjust()*.

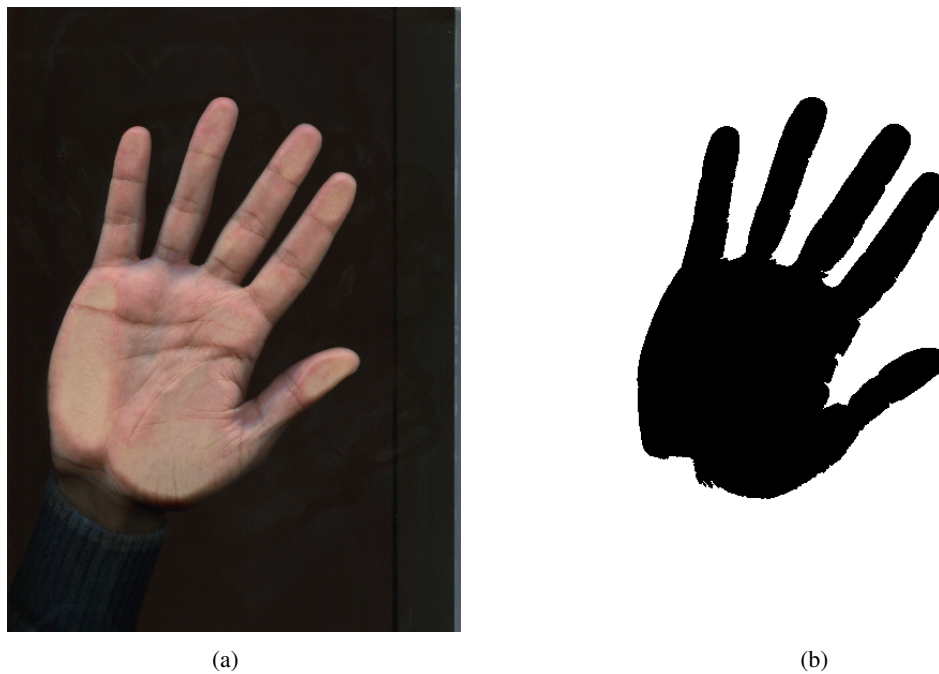


Figura 4.16: a) Imagem original RGB; b) Respectiva máscara binária mostrando os bons resultados do método de *Otsu*.

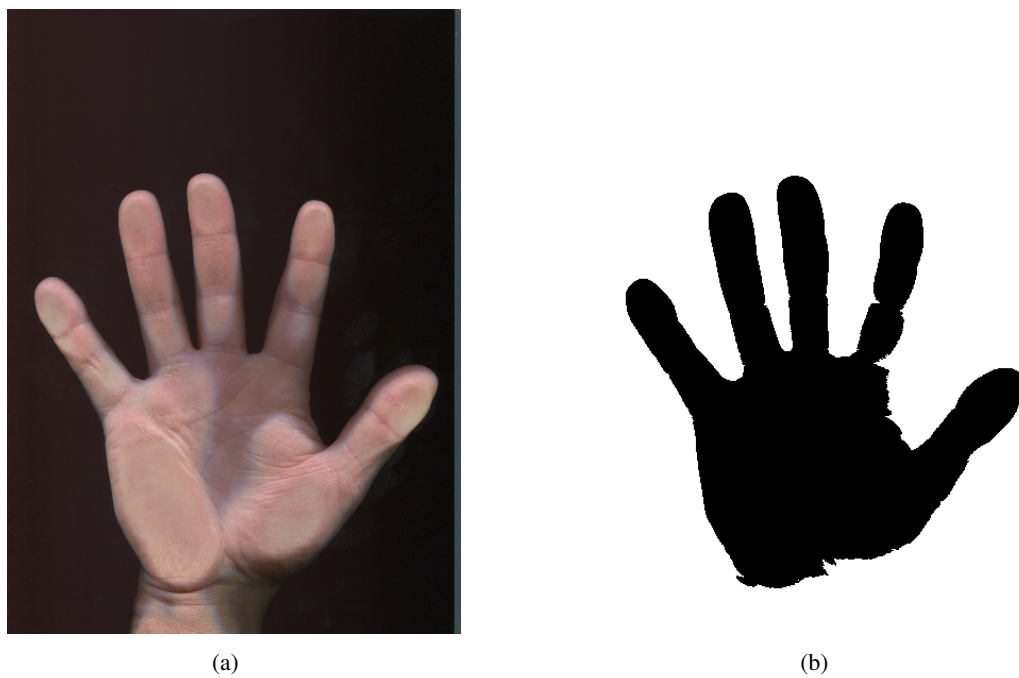


Figura 4.17: a) Imagem original RGB com presença de sombras; b) Respectiva máscara binária mostrando os bons resultados do método de *Otsu* para este caso.

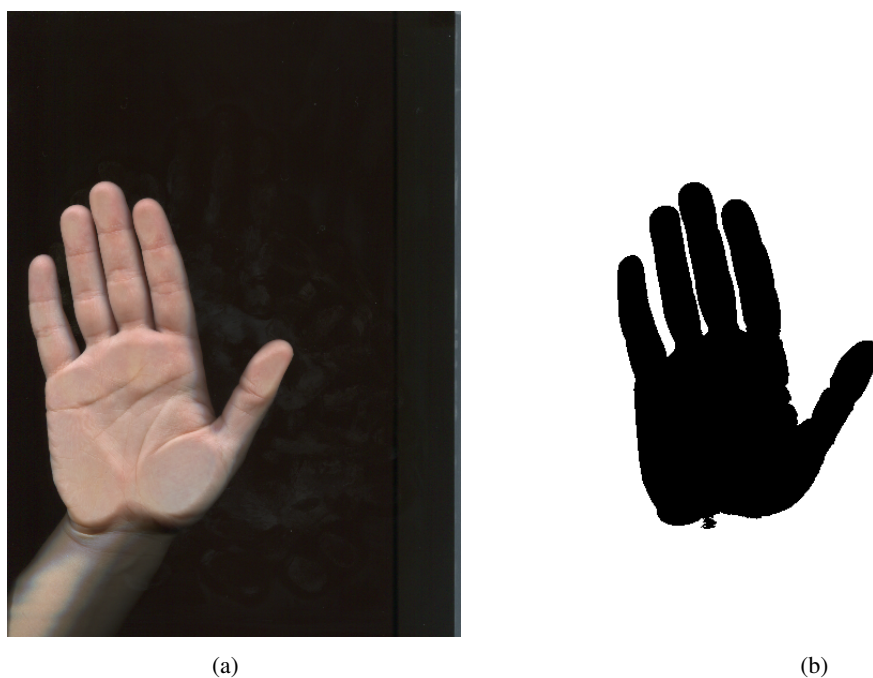


Figura 4.18: a) Imagem original RGB com dedos juntos; b) Respectiva máscara binária mostrando os bons resultados do método de *Otsu* para este caso.

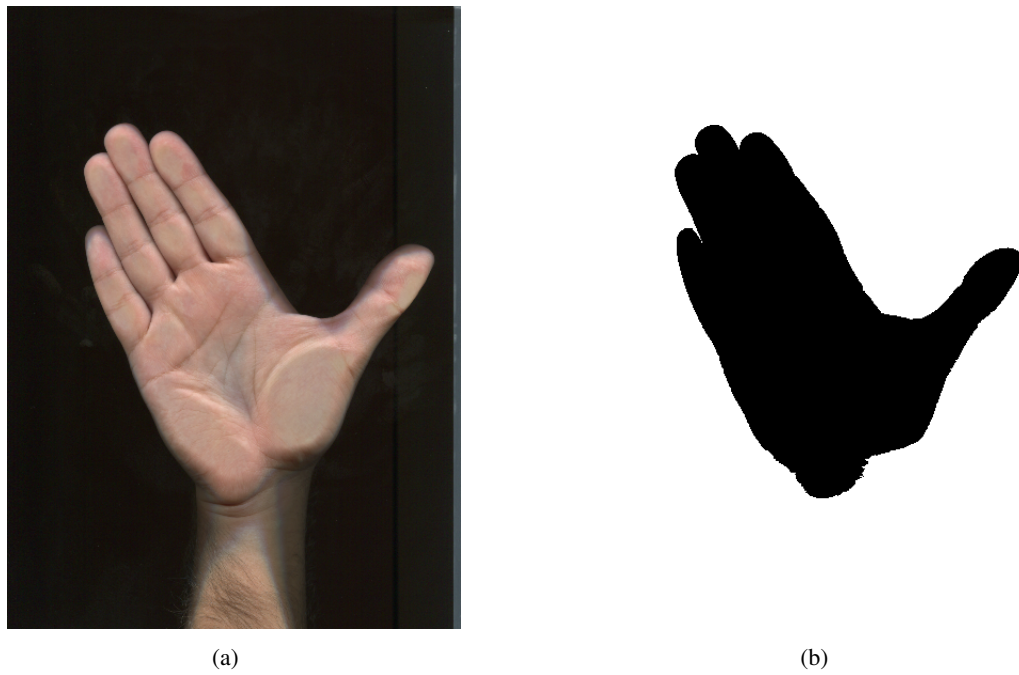


Figura 4.19: a) Imagem original RGB com dedos completamente juntos; b) Respectiva máscara binária mostrando a incapacidade de uma correcta segmentação da mão usando o método de *Otsu*.

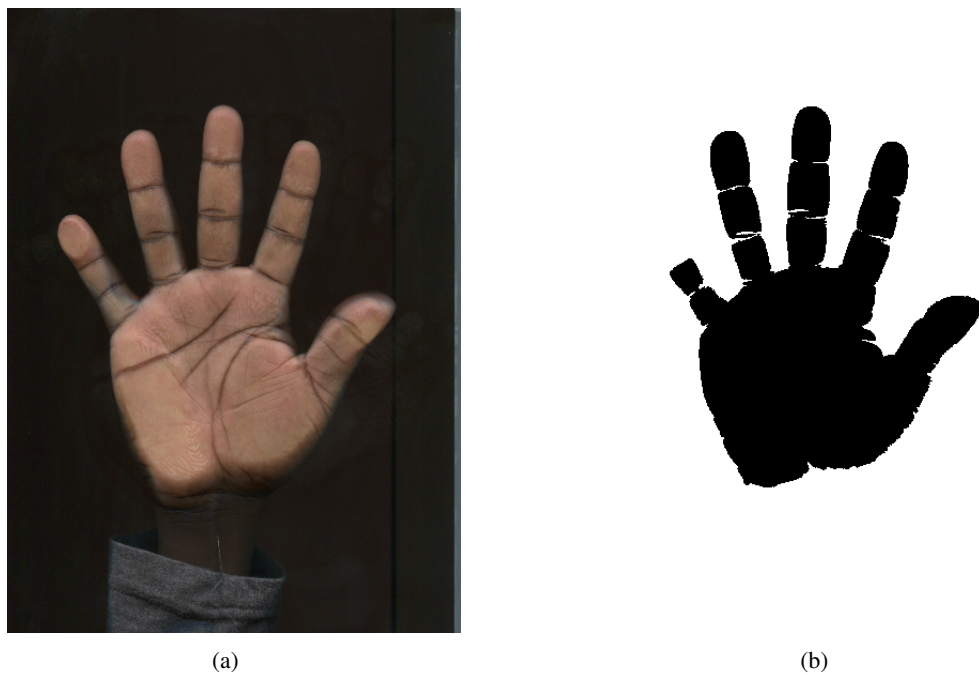


Figura 4.20: a) Imagem original RGB de individuo de raça negra; b) Respectiva máscara binária usando o método de *Otsu*, omitindo algumas zonas da mão devido a uma má segmentação.

Método IV - Ajuste Manual do Valor de *Threshold*

Dado todos os métodos anteriores não terem alcançado o sucesso total na binarização das imagens, seja em situações em que os dedos estão juntos, seja em casos em que existem pequenas sombras ou em que os sujeitos são de raças que não a caucasiana, foi feito um estudo do nível óptimo do *threshold* a aplicar em cada uma das imagens. O objectivo deste estudo é descobrir se existe um *threshold* óptimo para cada uma das imagens e que efectue correctamente a binarização para posteriormente, implementar uma função dedicada à binarização de imagens da mão, específica para este tipo de sistemas.

Deste modo, foram testadas várias imagens, principalmente aquelas que apresentavam maior taxa de insucesso noutros métodos, para determinar se existe sempre um limiar que separa correctamente a região da mão do fundo da imagem.

Resultados

Após testadas todas as imagens verificou-se que, em alguns casos, onde os métodos anteriores não apresentavam bons resultados, existe um nível de *threshold* que possibilita uma correcta binarização da mão (ver Figura 4.21 e Figura 4.22). Existem, no entanto, algumas imagens em que não é possível definir um limiar de binarização que permita uma correcta segmentação da zona da mão sem excluir partes desta e que separe devidamente os dedos. Nestes casos, ou se consegue estabelecer um nível de *threshold* que garanta a conectividade entre todas as partes da mão, mas que ao mesmo tempo não separa correctamente os dedos entre si fazendo com que, posteriormente, o contorno seja mal determinado, (Figura 4.23(b) e Figura 4.24(c)), ou então é estabelecido um *threshold* que permite separar totalmente os dedos mas que causa a perda da conectividade entre algumas partes (dedos), (Figura 4.23(c) e Figura 4.24(c)).

A partir da análise dos resultados aqui apresentados rapidamente se chega à conclusão de que não existe um valor para o *threshold* que permita uma binarização para todas as imagens da BD nas condições necessárias para o restante processamento a ser aplicado.

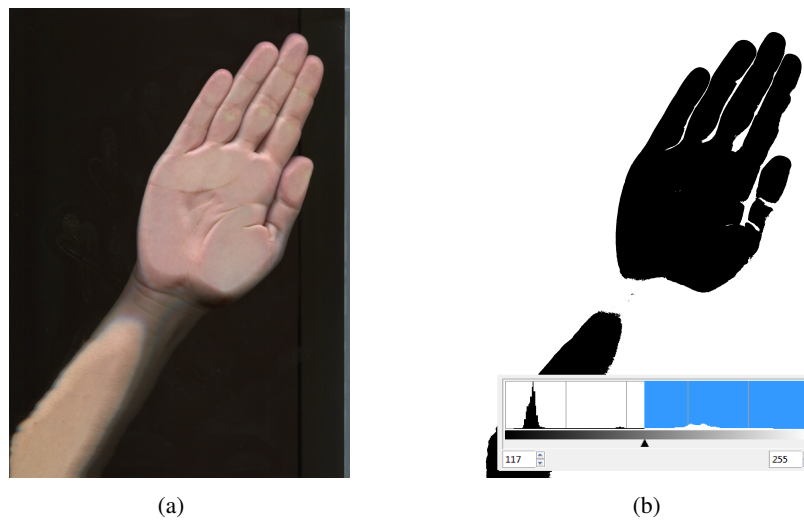


Figura 4.21: a) Imagem original RGB com dedos completamente juntos; b) Respectiva máscara binária mantendo conectividade entre todas as zonas da mão e, ao mesmo tempo,garantindo a separação completa entre os dedos. *Threshold* aplicado: 117/255.

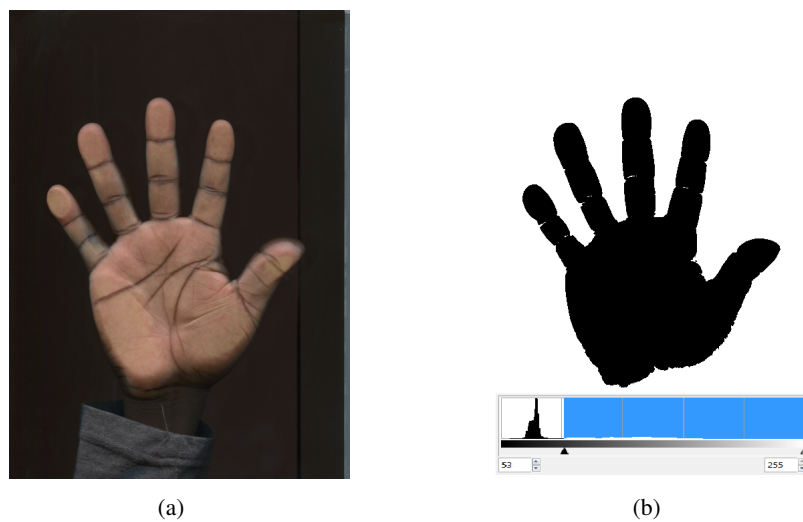


Figura 4.22: a) Imagem original RGB de individuo de raça negra; b) Respectiva máscara binária mantendo conectividade entre todas as zonas da mão e, ao mesmo tempo,garantindo a separação completa entre os dedos. *Threshold* aplicado: 53/255.

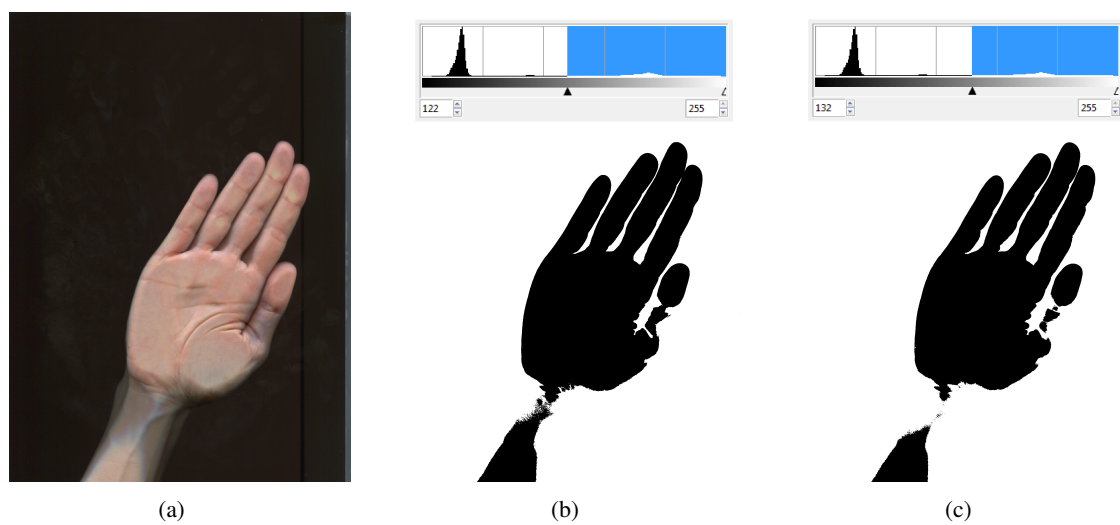


Figura 4.23: a) Imagem original RGB com dedos completamente juntos; b) Respectiva máscara binária mantendo conectividade entre todas as zonas da mão mas, ao mesmo tempo, não separando devidamente os dedos. *Threshold* aplicado: 122/255; c) Respectiva máscara binária garantindo uma correcta separação entre dedos mas não assegurando total conectividade entre todas as zonas da mão. *Threshold* aplicado: 132/255.

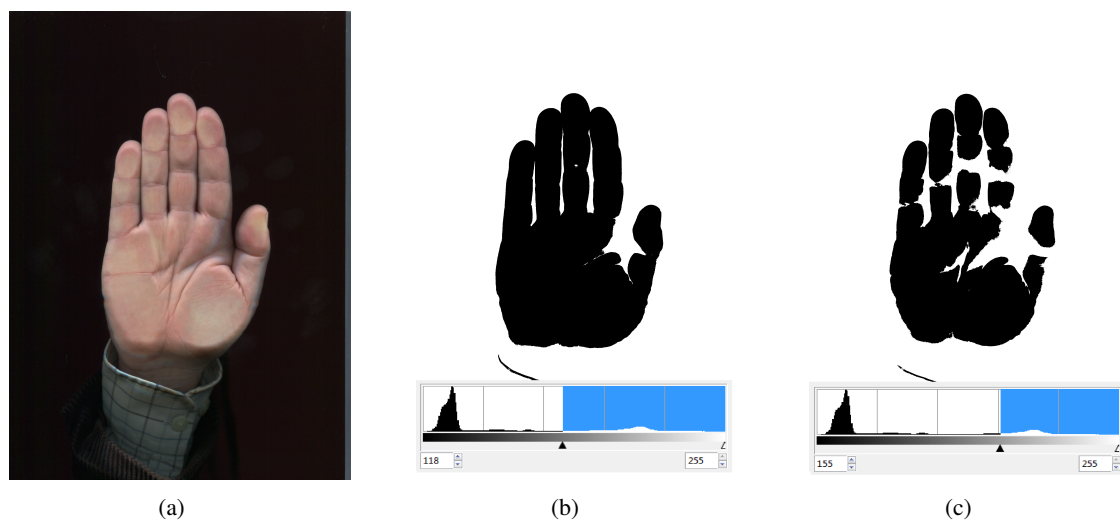


Figura 4.24: a) Imagem original RGB com dedos completamente juntos; b) Respectiva máscara binária mantendo conectividade entre todas as zonas da mão mas, ao mesmo tempo, não separando devidamente os dedos. *Threshold* aplicado: 118/255; c) Respectiva máscara binária garantindo uma correcta separação entre dedos mas não assegurando total conectividade entre todas as zonas da mão. *Threshold* aplicado: 155/255.

4.2.3 Sumário

A binarização da mão foi o primeiro passo da fase de pré-processamento e é responsável por identificar, isolar e extrair a mão do fundo da imagem, de forma a ser utilizada no próximo passo, a extracção de características. Foram aqui apresentados três métodos de binarização e nenhum deles se mostrou capaz de funcionar em todas as imagens da BD. A maioria dos métodos tende a falhar principalmente nas imagens que apresentam a mão com os dedos totalmente juntos e nas imagens de indivíduos de raça negra. Posto isto, surgiu a necessidade de escolher um método para se poder avançar com o desenvolvimento da fase de pré-processamento. De entre os três métodos propostos aquele que apresentou uma melhor relação entre desempenho e menor número de falhas foi o método III - Método de Otsu. Este método mostra-se muito rápido, estável e apresenta a maior taxa de eficácia na binarização quando comparado com os restantes métodos aqui testados. Será então o método a usar de aqui em diante.

4.3 Extracção de Contornos

4.3.1 Introdução

Identificada a mão na imagem original, é agora vez de extrair o contorno da mesma. No passo anterior, a segmentação binária, são identificados todos os objectos resultantes dessa operação e é calculada a área correspondente a cada objecto, sendo que o objecto que apresenta a maior área corresponde à mão. É então criada uma nova imagem binária contendo apenas o objecto de maior área e é garantido que este objecto não tem nenhum buraco (preenchendo com valor zero todos os espaços a branco (1) dentro do objecto-pai, *imfill()*). Assim, é garantido que a imagem que serve como parâmetro de entrada no passo de extracção de contornos contém apenas um objecto, sem buracos, que corresponde à mão.

4.3.2 Métodos e Resultados

Método I - Algoritmo de *Moore Neighbor*

Tendo uma imagem binária com um objecto, torna-se bastante simples a extracção do contorno do mesmo. Aqui será apresentado o algoritmo de *Moore Neighbor* com o critério de paragem de *Jacob's* (Pradhan et al., 2010). O algoritmo original de *Moore* está definido para contornos de vários pixels de largura. Requer visitar o pixel inicial duas vezes ou utilizar critérios de paragem

como o de Jacob. Neste algoritmo, a ideia básica é a seguinte: quando o pixel actual é preto(zero), a "vizinhança" *Moore* de P é examinada no sentido horário até que não seja encontrado nenhum pixel preto. De seguida, passamos para o pixel inicial e a vizinhança *Moore* de P é examinada agora no sentido anti-horário até que não reste nenhum pixel preto. O algoritmo de Moore-Neighbor encontra-se descrito abaixo:

Algoritmo *Moore Neighbor com critério de paragem de Jacob's*

Input: Imagem binária T contendo um objecto conectado de pixels pretos(zero) e fundo branco(um)

Output: Vector B contendo o pixels pertencentes ao contorno do objecto.

- Da esquerda para a direita e de cima para baixo, é feita uma procura na matriz da imagem até ser encontrado o primeiro pixel preto (zero) p ;
- Seja $M(p)$ a vizinhança *Moore* do pixel p . c representa o pixel actual sob consideração, ou seja, c pertence a $M(p)$;
- A vizinhança $M(p)$ é analisada no sentido horário e quando o pixel sob consideração, c , valer zero é inserido em B e $p=c$;
- O algoritmo termina quando o pixel inicial for visitado novamente.

Resultados

Tal como pode ser constatado na Figura 4.25 e Figura 4.26, este método produz bons resultados na extracção do contorno, desde que a imagem se encontre nas condições indicadas para tal, pois o sucesso deste passo depende totalmente dos resultados do passo anterior. Se a imagem proveniente da segmentação contém mais do que um objecto ou o objecto correspondente à mão não está devidamente demarcado (pontas dos dedos juntas), então haverá uma deficiente extracção de contornos.

Método II - Método de *Canny* para Detecção de Bordas

Ainda que o método apresentado anteriormente exiba excelentes resultados, ele tem a desvantagem de que se a imagem não estiver nas condições devidas, os resultados não serão os desejados. Posto isto, será testado um outro método para detecção do contorno, o método de *Canny*.

O detector de bordas de *Canny* é um operador Gaussiano de primeira ordem que atenua o ruído e localiza as bordas numa imagem.

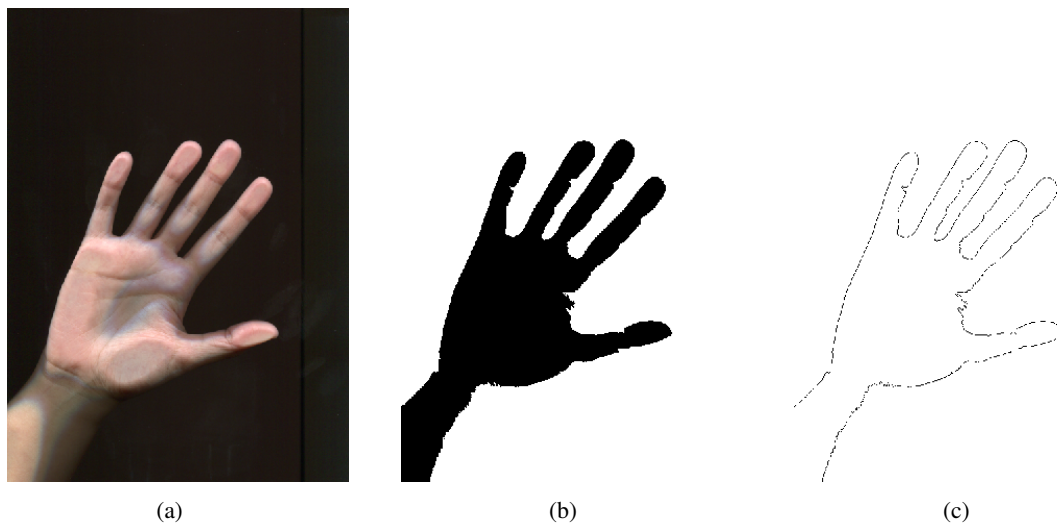


Figura 4.25: a) Imagem original RGB; b) Respectiva máscara binária usando o método de Otsu; c) Contorno da mão usando o método de *Moore Neighbor*.

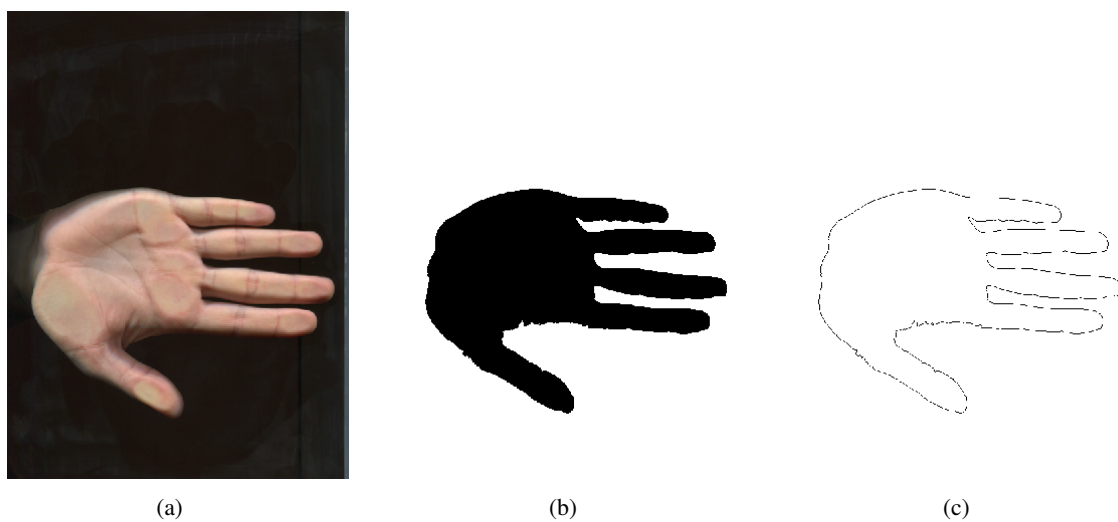


Figura 4.26: a) Imagem original RGB; b) Respectiva máscara binária usando o método de Otsu; c) Contorno da mão usando o método de *Moore Neighbor*.

Este algoritmo baseia-se nas seguintes regras:

- **Boa detecção:** deve possuir habilidade para localizar e marcar todas as bordas realmente existentes.
- **Boa localização:** minimizar a distância entre a borda detectada e a borda verdadeira (real).

- **Boa resposta:** para cada borda deve haver somente uma resposta.

Para satisfazer estas regras, o processo de detecção de bordas baseia-se em quatro passos:

Passo 1 - Normalização da imagem A imagem é normalizada por uma função Gaussiana bidimensional (2D) de tamanho especificado por um parâmetro comum. Na prática, convoluções Gaussianas 2D de tamanho elevado levam muito tempo a serem processadas, portanto é comum aproximá-las por duas funções Gaussianas unidimensionais, uma no eixo dos xx e a outra no eixo dos yy . Isso resulta em dois valores para cada pixel.

Passo 2 - Diferenciação Considerando a convolução 2D apresentada no passo anterior, a imagem normalizada é separada nas direcções segundo xx e segundo yy . Assim é possível calcular o gradiente da superfície uniforme da imagem convolucionada. Considerando também a aproximação unidimensional apresentada no Passo 1, os valores normalizados na direcção x são convolucionados utilizando a primeira derivada da função Gaussiana unidimensional com o mesmo alinhamento na direcção y . Da mesma maneira, os valores normalizados na direcção y são convolucionados utilizando a primeira derivada da função Gaussiana unidimensional com o mesmo alinhamento, agora na direcção x . Para somar os valores dos gradientes de x e y , a magnitude e o ângulo da inclinação podem ser calculados através da hipotenusa e $arctg$ do ângulo, de uma forma semelhante ao operador de *Sobel*.

Passo 3 - Omissão de pontos de mínima intensidade Encontrada a medida da intensidade de cada ponto da imagem, é agora necessário localizar as bordas. Isto é possível através da localização dos pontos de máxima intensidade ou, de maneira inversa, pelos pontos de mínima intensidade, que precisam de ser omitidos. Um valor de máxima intensidade ocorre no local mais alto da função de gradiente ou onde a derivada da função gradiente possui valor zero. Entretanto, desejamos omitir os pontos de mínima intensidade ou pontos de mínimas direcções perpendiculares à borda.

Passo 4 - Limiarização da borda (*threshold*) A limiarização usada no algoritmo de *Canny* usa o método denominado por "histerese". Considerando um segmento de borda para todo o valor situado acima do limite superior de limiarização, este é imediatamente aceite. Para todo o valor situado abaixo do limite inferior de limiarização, este é imediatamente rejeitado. Pontos situados entre os dois limites serão aceites se estiverem relacionados com pixels que apresentem respostas fortes.

Resultados

De forma a tentar contornar o problema acima descrito, no qual a binarização não tem o sucesso esperado e a imagem não está nas condições ideais para extracção de bordas, recorreu-se ao método de detecção de bordas de *Canny*. Com este método tentou extrair-se, directamente da imagem em tons de cinzento, o contorno desejado. O objectivo da aplicação deste método é analisar os seus resultados para todas as imagens da BD, principalmente aquelas em que outros métodos de binarização tendiam a falhar.

A função utilizada para a implementação do algoritmo de *Canny* pertence a [Kovesi](#) e utiliza como parâmetros de entrada a imagem em escala de cinzentos e o valor de σ do filtro de suavização Gaussiano. Como saídas, esta função devolve uma imagem com a orientação e uma outra com a força das bordas presentes na imagem (imagem gradiente), Figura 4.27.

Após análise deste primeiro resultado ficou claro que seria necessário algum processamento na mesma para alcançar os resultados desejados. Assim sendo, foi aplicado um *threshold* na imagem gradiente de forma a eliminar as bordas mais fracas e a manter as mais fortes, (Figura 4.28(a)).

Foram então aplicadas algumas operações morfológicas para diminuir o ruído presente na imagem. Primeiro foi aplicada uma operação de esqueletização na imagem da Figura 4.28(a) para reduzir o contorno da mão a uma linha com espessura de um pixel apenas. O resultado, embora apresente melhorias significativas, contém ainda alguns segmentos do esqueleto que devem ser removidos. Isso foi conseguido recorrendo à operação de "poda" do esqueleto. No entanto, como não é possível saber para cada caso qual o parâmetro a introduzir nesta operação (limite de "poda"), existem casos em que é demasiado extensa e existem outros em que não é suficiente para eliminar todos os segmentos de esqueleto indesejados. Por defeito utilizou-se "poda" infinita, garantindo assim que o esqueleto fosse reduzido a uma linha apenas, mas não sendo garantido que o início e fim do esqueleto coincidisse com o início e fim reais do contorno, tal como pode ser verificado na Figura 4.28(c).

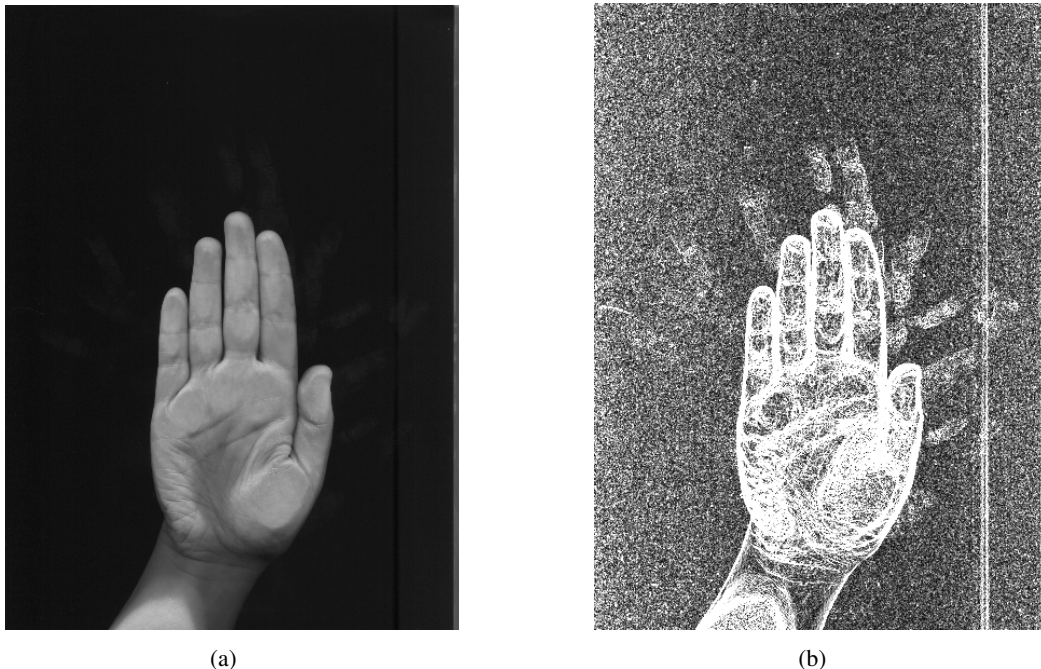


Figura 4.27: a) Imagem original em tons de cinzento; b) Respectiva imagem gradiente, resultado da aplicação do algoritmo de *Canny*.

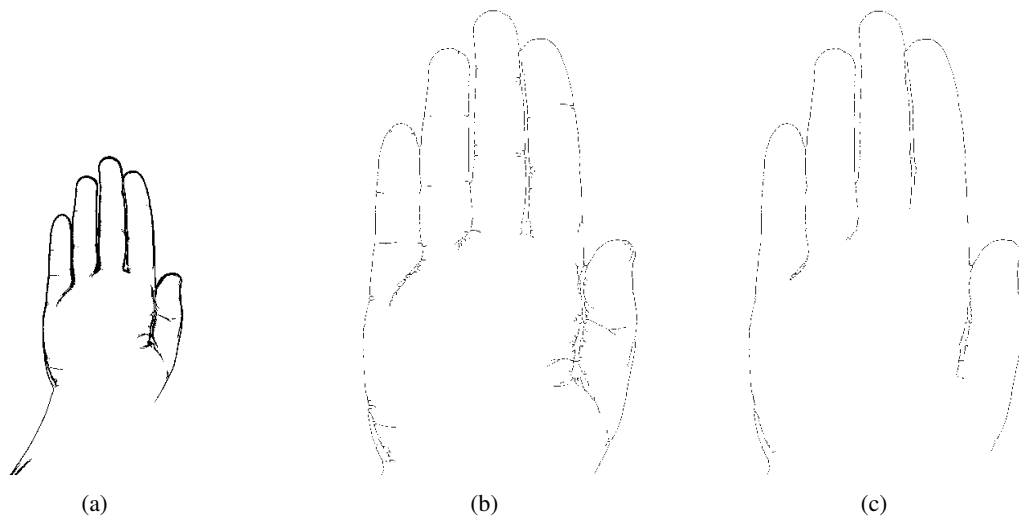


Figura 4.28: a) Imagem gradiente de *Canny* depois de aplicado o *threshold* para manter apenas as bordas mais fortes; b) Respetivo esqueleto (resultado da aplicação da operação morfológica de esqueletização); c) Esqueleto final, após aplicação da operação morfológica de poda.

4.3.3 Sumário

Foi apresentado um método de extracção de contornos, o método de *Moore Neighbor*, que oferece excelentes resultados. No entanto, este método impõe a condição de que a imagem a ser tratada tem de cumprir certas regras: são elas na imagem apenas pode estar presente um objecto; esse objecto tem de corresponder à mão e tem de estar correctamente demarcado.

Apesar do esforço feito no passo da segmentação para que todas as imagens cumprissem tais regras, existem algumas em que isso não é possível. Na medida de combater estes casos de insucesso foi também aqui proposta a implementação de métodos de extracção do contorno directamente da imagem original.

O método aqui apresentado foi o método de detecção de bordas de *Canny*. Este método é capaz de identificar todas as bordas para as imagens em que os dedos se encontram totalmente juntos (as que geralmente falham na binarização), mas, apesar de realizadas várias operações morfológicas para a extracção do contorno como uma linha única, fechada e com espessura de um pixel, tal não foi possível.

O resultado de todo o conjunto de operações feitas nesse sentido resulta sempre num contorno não fechado e inconstante, ou seja, o seu início e o seu fim nunca podem ser correctamente identificados e a zona entre os dedos também não se encontra em condições de ser tratada como um contorno.

Outros filtros, semelhantes ao de *Canny*, foram aqui igualmente testados (Sobel, Prewitt, etc.). Contudo, estes produziram resultados semelhantes aos de *Canny*: nunca apresentaram um resultado final melhor. Dados estes resultados, o método a ser aplicado nas fases seguintes será o inicialmente descrito, *Moore Neighbor*.

4.4 Detecção do ângulo de orientação e alinhamento das imagens

4.4.1 Introdução

Cumpridos os passos de obtenção da máscara binária e respectiva extração de contornos, é necessário alinhar a imagem da mão para que os métodos desenvolvidos sejam mais facilmente aplicados. Este passo completa a fase de pré-processamento e consiste em determinar o ângulo de inclinação da mão e com essa informação rodar a imagem. Esta etapa é de grande importância pois, no fim, permite ao sistema ter algum conhecimento sobre a disposição da mão. Quando se utiliza uma BD como esta, em que existem imagens com total liberdade de rotação, o correcto alinhamento é crucial. Por exemplo, nos casos em que a mão se encontra invertida 180° ou 90°, se não for efectuado um alinhamento prévio, os algoritmos a desenvolver nas fases seguintes serão mais complexos.

Neste passo o alinhamento pode ser feito nas imagens binarizadas ou nas imagens do contorno. As principais diferenças residem em processar algumas centenas de pixels, caso do contorno, ou alguns milhares deles, caso da máscara binária. Constatou-se também que o cálculo do Centro de Massa (CM) é mais preciso quando feito em imagens da máscara binária. Contudo, dependendo dos algoritmos desenvolvidos nas fases posteriores, o alinhamento poderá ser aplicado a ambos os tipos de imagens (máscara binária ou contorno).

Existem diversas formas de efectuar o alinhamento da imagem. Nos trabalhos de [Wong e Shi \(2002\)](#), [Adan et al. \(2008\)](#) e [Kumar et al. \(2003\)](#) é apresentada uma forma que consiste em definir o eixo principal do dedo médio como o eixo vertical do sistema de referência (yy), (Figura 4.29). Contudo, para que isso aconteça, é necessário que sejam primeiramente identificados os pontos correspondentes à ponta do dedo médio e à sua base.

Desta forma decidiu-se optar por um outro método de alinhamento. Em primeiro lugar, ainda não foram detectados quaisquer pontos característicos para a possibilidade de realizar um alinhamento como o referido anteriormente. Em segundo lugar, devido às inúmeras configurações em que os dedos podem ser colocados, um alinhamento a partir da posição de qualquer um dos dedos não seria consistente para toda a BD.

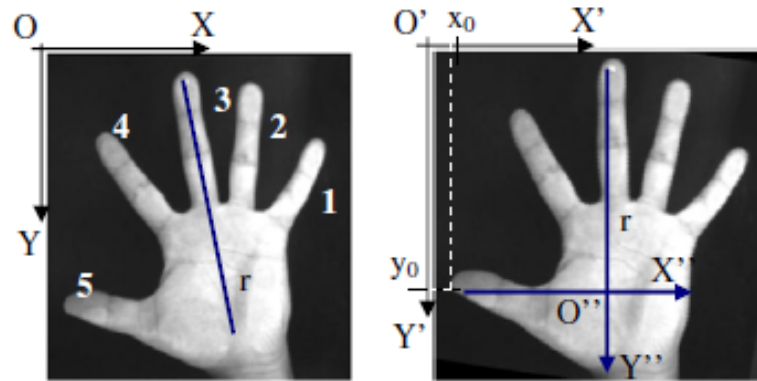


Figura 4.29: Método de alinhamento proposto por Wong e Shi (2002) e Adan et al. (2008).

4.4.2 Métodos e Resultados

Método I - Método dos Momentos

Para o alinhamento de uma imagem é necessário saber, em primeiro lugar, o centro de massa (X_{cm}, Y_{cm}) do objecto e, em segundo lugar, ter conhecimento do ângulo θ de orientação do mesmo. O alinhamento consiste em rodar o objecto θ graus em torno do respectivo CM, de forma a que o eixo principal do objecto coincida com o eixo dos yy no referencial final. O método dos momentos foi o escolhido para este passo importante, uma vez que permite obter toda a informação necessária. A definição de momento de um objecto na imagem $I(r,c)$ é a seguinte:

$$m_{p,q} = \sum_{x=1}^c \sum_{y=1}^r x^p \cdot y^q \cdot I(x,y) \quad (4.4)$$

onde:

- I representa a imagem com r linhas e c colunas;
- $I(x,y)$ representa o pixel na posição (x,y) da imagem e vale um se pertence ao objecto, ou zero se não pertence.

Os momentos são geralmente indicados pela sua ordem. Esta ordem depende do índice \mathbf{p} e \mathbf{q} , representado pela soma de $\mathbf{p+q}$. Considerado isto, os momentos até à segunda ordem são definidos como:

- Momento de ordem zero:

$$m_{0,0} = \sum_{x=1}^c \sum_{y=1}^r I(x,y) \quad (4.5)$$

- Momento de primeira ordem:

$$m_{1,0} = \sum_{x=1}^c \sum_{y=1}^r x \cdot I(x,y) \quad (4.6)$$

$$m_{0,1} = \sum_{x=1}^c \sum_{y=1}^r y \cdot I(x,y) \quad (4.7)$$

- Momento de segunda ordem:

$$m_{2,0} = \sum_{x=1}^c \sum_{y=1}^r x^2 \cdot I(x,y) \quad (4.8)$$

$$m_{0,2} = \sum_{x=1}^c \sum_{y=1}^r y^2 \cdot I(x,y) \quad (4.9)$$

$$m_{1,1} = \sum_{x=1}^c \sum_{y=1}^r x \cdot y \cdot I(x,y) \quad (4.10)$$

Os momentos representam características de um objecto. Não têm um significado geométrico, mas normalmente, podem ser derivados deles parâmetros geométricos:

- O momento de ordem zero $m_{0,0}$ representa a área do objecto;
- O CM (X_{cm}, Y_{cm}) do objecto é descrito pelos momentos de primeira ordem $m_{1,0}$ e $m_{0,1}$ divididos pelo momento de ordem zero $m_{0,0}$;
- A orientação θ do objecto é definida entre o eixo do xx e a direcção principal do objecto, Figura 4.30.

O cálculo do ângulo θ é demonstrado de seguida:

$$\theta = \frac{1}{2} \cdot \tan^{-1} \cdot \frac{2 \cdot m_{1,1}}{m_{0,2} - m_{2,0}} \quad (4.11)$$

Contudo, esta não é a equação genérica para todos os ângulos de orientação, (ver Tabela 4.1 para observar todos os casos possíveis).

Após a definição da inclinação e do CM do objecto é possível fazer corresponder o eixo principal do objecto ao eixo vertical do sistema de coordenadas final. Para cada uma das imagens foi aplicada uma matriz de rotação que, a partir da informação de θ e das coordenadas originais da

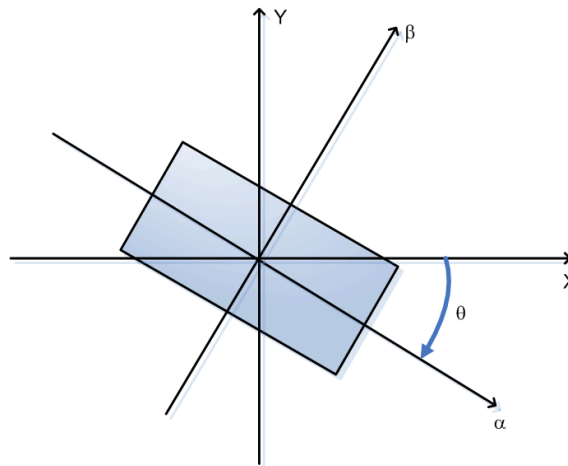


Figura 4.30: Sistemas de referência do objecto e final.

Tabela 4.1: Diferentes casos no cálculo da orientação.

$m_{2,0} - m_{0,2}$	$m_{1,1}$	θ	θ range
Positivo	0	0	
Negativo	0	π	
Negativo	Positivo	$\frac{1}{2} \cdot \tan^{-1} \cdot \frac{m_{2,0} - m_{0,2}}{2 \cdot m_{1,1}} - \frac{\pi}{4}$	$\frac{\pi}{4} < \theta < \frac{\pi}{2}$
Negativo	Negativo	$\frac{1}{2} \cdot \tan^{-1} \cdot \frac{m_{2,0} - m_{0,2}}{2 \cdot m_{1,1}} + \frac{\pi}{4}$	$-\frac{\pi}{2} < \theta < -\frac{\pi}{4}$
Positivo	$\neq 0$	$\frac{1}{2} \cdot \tan^{-1} \cdot \frac{2 \cdot m_{1,1}}{m_{0,2} - m_{2,0}}$	$-\frac{\pi}{4} < \theta < \frac{\pi}{4}$

mão, tornou possível a obtenção das coordenadas segundo a nova base do referencial. A matriz de rotação iR_j apresenta o seguinte aspecto:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \cos(\theta(t)) & -\sin(\theta(t)) \\ \sin(\theta(t)) & \cos(\theta(t)) \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \quad (4.12)$$

onde, α e β representam os eixos do novo sistema de coordenadas de referencia. x e y representam o referencial original.

Obtidos todos os parâmetros necessários à rotação, partiu-se então para a implementação do algoritmo. A função de rotação num primeiro passo determina o CM do objecto, de seguida calcula o ângulo θ de orientação e, por fim, é aplicada a matriz de rotação apresentada acima. No entanto para a aplicação da referida matriz de rotação são necessários três passos:

1. Transladar o objecto de forma a fazer coincidir o seu CM com a origem da imagem (coordenadas (0,0));
2. Aplicar a matriz de rotação iR_j ;

3. Transladar de novo o objecto e fazer coincidir o seu CM com o centro da imagem.

Logo se percebeu que a implementação destes passos iria levar a uma diminuição do desempenho do sistema, pois estas operações tornavam-se extremamente pesadas, principalmente quando se testava toda a BD. Assim sendo, foi utilizada uma função disponibilizada pelo MatLab, *imrotate()*, que efectua a rotação de objectos apenas a partir do ângulo θ . Esta função mostrou-se bastante mais rápida e apresenta igualmente bons resultados. Apenas no final desta operação de rotação é realizada uma translação do objecto, de forma a centrá-lo na imagem, isto facilitará o processamento nos passos seguintes.

Imagens com a mão invertida

Existem casos em que a mão se encontra invertida, isto é, com os dedos na parte inferior. Nestes casos o método dos momentos apenas determina o ângulo de inclinação do objecto (mão), não detectando, no entanto, se esta se encontra inversamente posicionada ou não. Para tal, foi implementado um método de detecção da sentido da mão que apresenta o seguinte funcionamento:

- Efectuar a rotação utilizando a função *imrotate()*;
- Fazer dois varrimentos, V1 e V2, na imagem alinhada, V1 acima da linha do CM e V2 abaixo dessa mesma linha;
 - SE V1 detecta transições (preto->branco) correspondentes aos 4 dedos, ENTÃO:
 - * a mão está correctamente alinhada. Fim do alinhamento.
 - SE V2 detecta transições (preto->branco) correspondentes aos 4 dedos, ENTÃO:
 - * a mão está invertida.
 - * efectuar nova rotação de 180° . Fim do alinhamento.

4.4.3 Resultados

Embora a implementação do algoritmo de rotação recorrendo à matriz iR_j tenha colocado alguns obstáculos, estes foram ultrapassados tendo-se obtido resultados bastante satisfatórios. Contudo, optou-se por usar a função de rotação já disponível no MatLab com a vantagem de simplificar o algoritmo descrito anteriormente. Os resultados, (Figura 4.31) apresentados por esta função correspondem integralmente aos esperados, não existindo assim qualquer ponto negativo a apontar neste passo de alinhamento da imagem. Mesmo em casos onde a mão se encontra invertida ou na

horizontal, o alinhamento é feito de forma correcta e apresenta sempre como resultado a imagem com a mão alinhada na vertical, o que pode ser comprovado na Figura 4.32 e Figura 4.33.

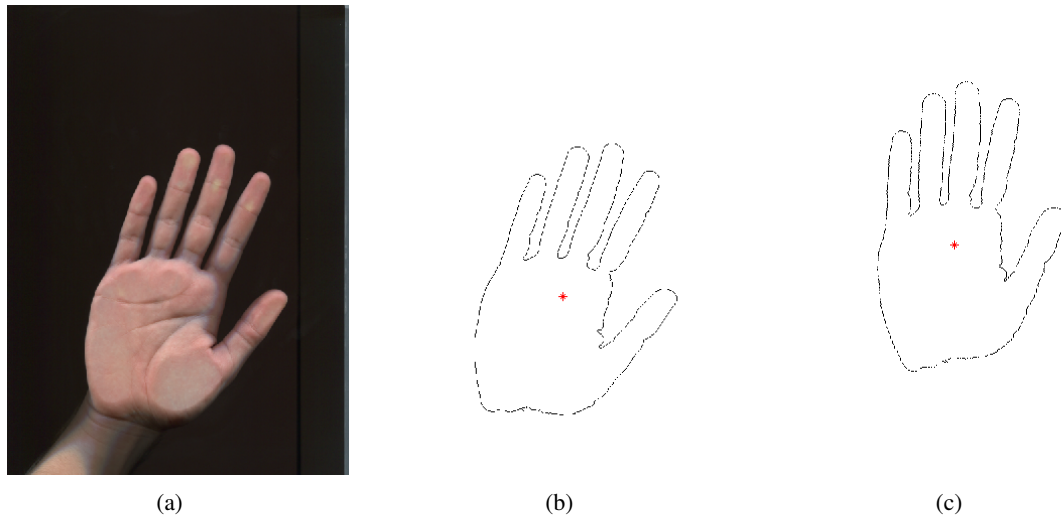


Figura 4.31: a) Imagem original RGB; b) Respectiva imagem do contorno no sistema referencial inicial com o CM assinalado. Apresenta um ângulo de inclinação: $\theta = 75.06^\circ$, 1º Quadrante; c) Respectiva imagem do contorno alinhada com o sistema referencial final.

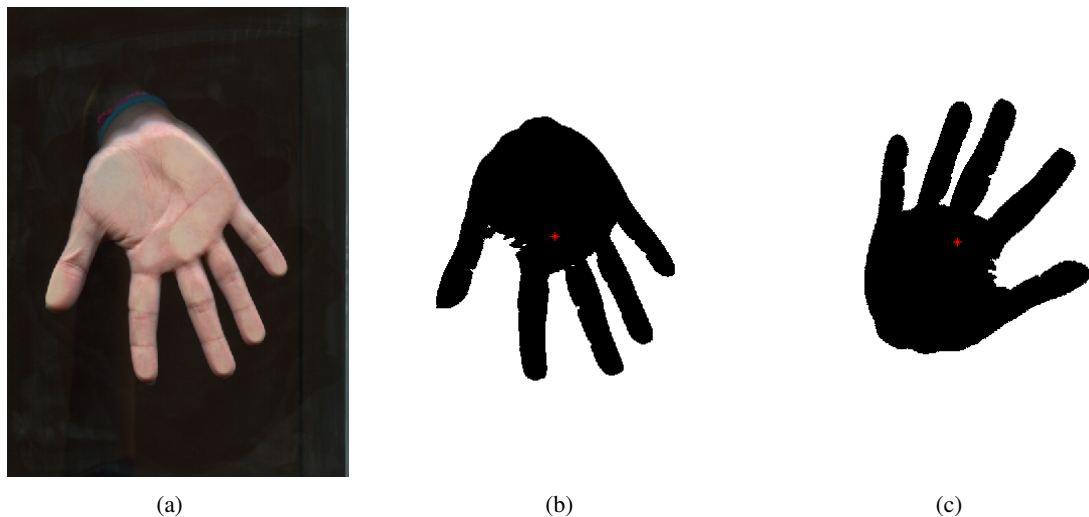


Figura 4.32: a) Imagem original RGB com a mão invertida; b) Respectiva imagem do contorno no sistema referencial inicial com o CM assinalado. Apresenta um ângulo de inclinação: $\theta = -46.53^\circ$, 4º Quadrante; c) Respectiva máscara binária alinhada com o sistema referencial final.

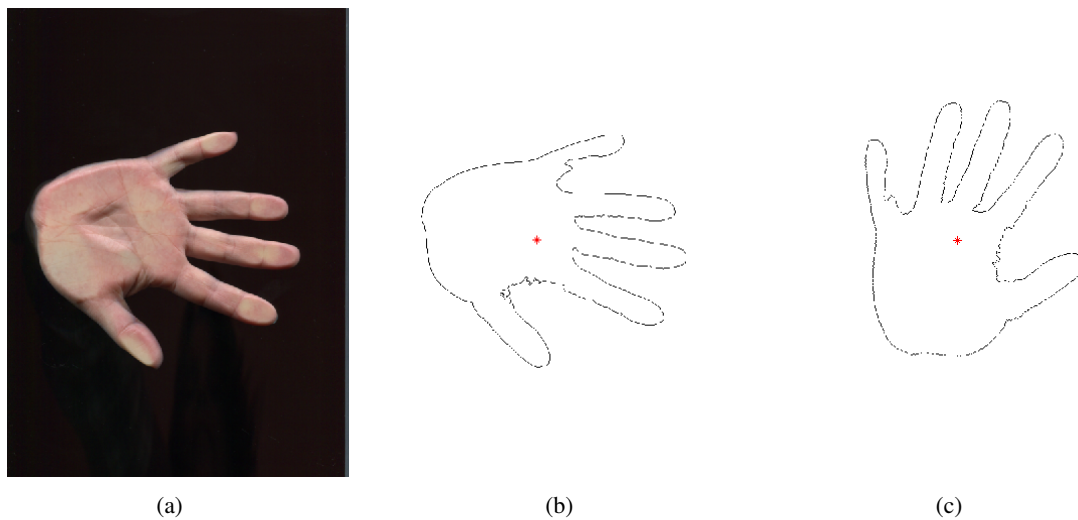


Figura 4.33: a) Imagem original RGB com a mão disposta na horizontal; b) Respectiva imagem do contorno no sistema referencial inicial com o CM assinalado. Apresenta um ângulo de inclinação: $\theta = 10.95^\circ$, 1º Quadrante; c) Respectiva imagem do contorno alinhada com o sistema referencial final.

4.4.4 Sumário

Foi aqui apresentado o método utilizado para o alinhamento das imagens da BD. Este passo tem como finalidade colocar todas as imagens nas mesmas condições e, desta forma, colocá-las em igualdade de circunstâncias. Este passo assume uma grande importância neste sistema devido ao facto de ser usada uma BD sem restrições na aquisição, ou seja, existe total liberdade de rotação na colocação da mão durante a aquisição. Como já foi verificado anteriormente, existem imagens na BD que apresentam os mais variados ângulos de orientação, desde 0° até 180° . Nestes casos, é essencial para o sucesso dos passos seguintes que todas estas imagens sejam alinhadas segundo o mesmo critério. Os resultados correspondem ao esperado tendo funcionado correctamente para todas as imagens da BD.

4.5 Conclusão

Foram aqui apresentadas as metodologias utilizadas durante a fase de pré-processamento. Esta fase foi constituída por três grandes passos, a segmentação, a extracção de contornos e, finalmente, o alinhamento da imagem.

Na primeira fase foi detectado um grande entrave para as imagens que apresentam os dedos juntos. Nestes casos, a binarização falha, não conseguindo separar os dedos. Apesar de testados vários métodos, o problema persistiu e foi então realizado um estudo acerca do nível de *threshold* óptimo de cada uma das imagens. Verificou-se, a partir deste estudo que nem todas as imagens possuíam um *threshold* que permitia a sua correcta binarização, o que levou à conclusão de que

métodos baseados num *threshold* único não funcionam para todas as imagens. No entanto, métodos baseados em *threshold* por zonas, como é o caso da função *Ridge-Segments()*, também não mostraram resultados optimistas.

Na extracção de contornos os resultados foram positivos. O contorno da mão é detectado correctamente em todas as imagens em que a binarização obteve o sucesso esperado. Assim, para todos aqueles casos em que a binarização falhou, consequentemente, a extracção do contorno falhou também. Para tentar combater este problema, foram testadas outras formas de obter directamente o contorno da mão sem recorrer à máscara binária. Os métodos de detecção de bordas de *Canny*, *Sobel* e *Prewitt* foram aqui testados e apesar de serem capazes de identificar o contorno em imagens em que os dedos se encontravam juntos, o contorno resultante não foi o desejável.

Finalmente, no alinhamento da imagem, o objectivo era alinhar a mão para que todas as imagens se encontrassem nas mesmas condições para as fases seguintes. Foi proposto o método dos momentos que permite detectar o CM da mão e o ângulo θ de inclinação. Assim a mão é alinhada de acordo com a sua distribuição de massa e todas as imagens são alinhadas de acordo com um parâmetro próprio.

Todos os métodos implementados nesta fase de pré-processamento tiveram em consideração que a BD possui apenas imagens da mão direita. Contudo, pequenas e simples alterações poderiam ser efectuadas de modo a tornar possível a sua aplicação em BD com imagens de mãos direitas e esquerdas.

A Figura 4.34 ilustra a passagem de uma imagem pela fase de pré-processamento:

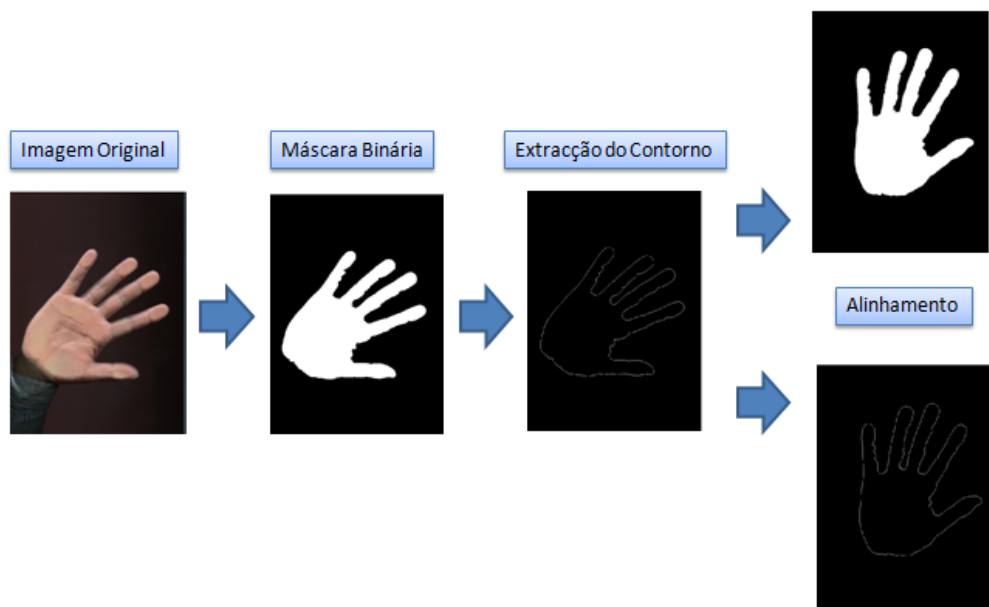


Figura 4.34: Passagem de uma imagem pela fase de pré-processamento.

Capítulo 5

Detecção de Pontos Característicos

5.1 Introdução

Concluída a fase de pré-processamento, segue-se a detecção dos pontos característicos da mão. Estes pontos são aqueles a partir dos quais é possível fazer a extracção de características. Segundo [Wong e Shi \(2002\)](#), [Faundez-Zanuy e Merida \(2005\)](#) ou [Montalvao et al. \(2010\)](#) existem nove pontos característicos da mão que permitem a extracção das características necessárias à fase de reconhecimento. Tais nove pontos correspondem às cinco pontas dos dedos e aos quatro vales entre dedos adjacentes, (Figura 5.1). Com estes pontos é possível determinar o comprimento, a largura dos dedos e algumas propriedades da região da ponta dos dedos. De entre as possíveis características a usar na fase da identificação, estes pontos permitem a extracção de uma grande parte delas. É por isso fundamental que a localização destes pontos seja a mais correcta e precisa possível, para desta forma proporcionar uma maior exactidão das medidas realizadas com base nestes pontos.

5.2 Métodos e Resultados

Aqui serão apresentados alguns métodos que foram implementados para a detecção dos 9 pontos. Alguns deles foram baseados nos existentes na literatura, outros foram implementados e testados pela primeira vez num sistema deste tipo.

5.2.1 Transformação para Coordenadas Polares e Análise da Curvatura

Este método de detecção dos pontos característicos da mão foi anteriormente usado em outros trabalhos ([Montalvao et al., 2010](#)) e consiste, basicamente, numa transformação de coordenadas. A imagem do contorno obtida na fase de pré-processamento encontra-se mapeada em coordenadas cartesianas e este método consiste em transformar tal imagem numa outra em coordenadas polares. Espera-se com isto facilitar a identificação dos pontos desejados.

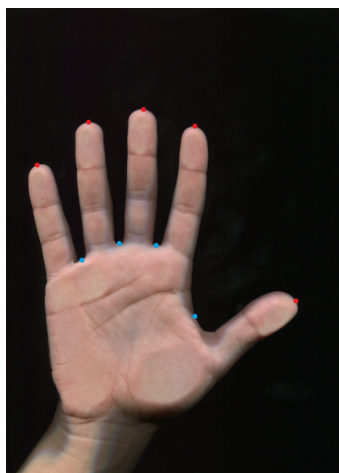


Figura 5.1: Representação dos 9 pontos característicos.

O primeiro passo foi então a transformação da imagem do contorno original numa imagem em coordenadas polares, o que foi realizado recorrendo a uma função específica que utilizou como centro para a transformação para coordenadas polares o CM da mão. No entanto, depois de algumas experiências, verificou-se que a parte inferior do contorno (pertencente a zonas do braço) introduzia muito ruído na imagem em coordenadas polares. Dado que a parte inferior da mão não contém nenhum tipo de informação relevante para extracção, decidiu-se eliminar a informação do contorno que se encontra abaixo do CM da mão. Tal operação introduziu melhorias significativas na imagem final em coordenadas polares, tal como pode ser verificado na Figura 5.2 e Figura 5.3.

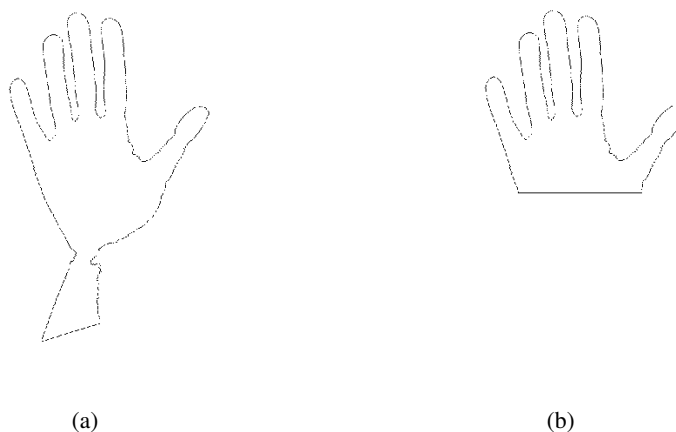


Figura 5.2: a) Imagem original alinhada do contorno; b) Imagem do contorno original cortada ao nível do CM.

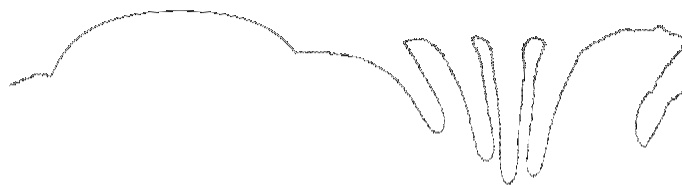


Figura 5.3: Imagem do contorno da parte superior da mão em coordenadas polares.

Após a obtenção da imagem do contorno em coordenadas polares, são visíveis as vantagens desta transformação para a análise e identificação dos pontos característicos.

Na etapa seguinte foi efectuada uma análise à curvatura nesta imagem. Para isso foi necessário implementar uma função de seguimento de contorno para extrair um vector com as coordenadas de cada ponto pertencente ao mesmo. Obtido esse vector foram mapeados todos os pontos do contorno num espaço cartesiano para uma análise mais fácil da primeira derivada do contorno. Recorrendo à análise da primeira derivada do contorno, é possível obter informação relativa à curvatura em cada ponto e assim procurar pelos nove pontos que apresentam o menor valor da referida característica. Contudo, verificou-se que o contorno contém ainda algum ruído, principalmente na zona dos vales entre os dedos, (Figura 5.4(a)). Para eliminar tal ruído, recorreu-se a uma suavização do contorno, aplicando-se um filtro de suavização existente na *toolbox* do MatLab (*smooth()*) utilizando o método *loess - Local regression using weighted linear least squares and a 2nd degree polynomial model*. Feita esta pequena operação de suavização e analisando a derivada do contorno, foi então possível obter resultados positivos na marcação dos pontos pela detecção de menor curvatura do contorno.

Resultados

Pela Figura 5.4(b) verifica-se que os pontos detectados pela análise à primeira derivada correspondem, de facto, aos nove pontos característicos. Identificados os pontos, foi efectuada um novo mapeamento do contorno no sistema de coordenadas cartesianas original, de forma a estes corresponderem aos pontos na imagem do contorno original. Este método, tal como pode ser constatado na Figura 5.5, produz resultados bastante satisfatórios.

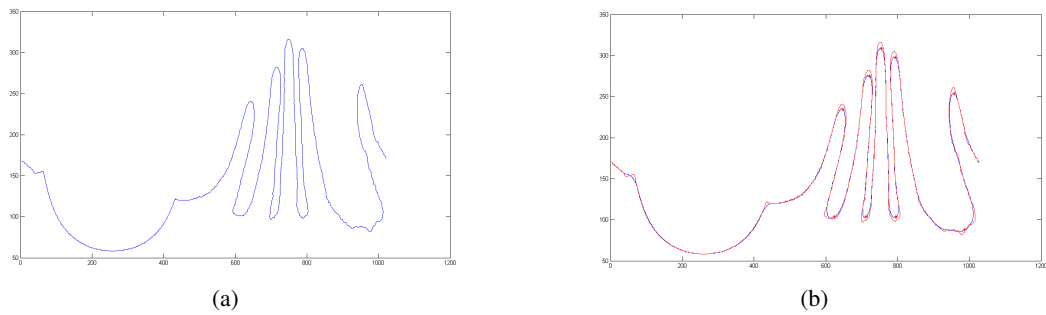


Figura 5.4: a) Pontos do contorno mapeados num espaço cartesiano; b) Comparação entre contorno original e contorno suavizado (com método *loess*), linha a vermelho - contorno original, linha a azul - contorno suavizado. Representação dos pontos identificados como os de maior curvatura, resultado da análise da primeira derivada.



Figura 5.5: Mapeamento dos pontos identificados na imagem original comprovando os bons resultados que este método apresenta.

Verificou-se, contudo, alguns resultados menos positivos com a utilização deste método. Por exemplo, o facto de trabalhar apenas com os pontos do contorno que se encontram acima da linha do CM pode, por vezes, introduzir alguns erros indesejados. Isto pode acontecer quando região superior à linha ao nível horizontal do CM não abrange os cinco dedos; isto é, quando o início do dedo polegar está abaixo do CM, não permanecendo por isso juntamente com a parte superior do contorno. Nestes casos, apenas segue o contorno de quatro dedos, (Figura 5.6 e Figura 5.7). No entanto, a pesquisa é feita da mesma forma, ou seja, são eleitos nove pontos com a maior

curvatura como os pontos característicos da mão. Dado que nestes casos apenas existem sete pontos característicos reais, significa que serão marcados dois pontos que não corresponderão a nenhum dos pontos característicos.

Existem, também, outros obstáculos encontrados durante os testes realizados neste método. Um deles tem origem na imagem do contorno em coordenadas polares. Isto é, após a transformação para coordenadas polares a imagem apresenta algum ruído à volta do contorno, e mesmo efectuando várias operações morfológicas, existem casos em que a imagem não apresenta as características óptimas para ser utilizada pela função de seguimento de contornos. Normalmente estes casos de insucesso devem-se à desconectividade entre dois pixels consecutivos, a um pixel ter conectividade com mais de 2 pixels ou quando o contorno apresenta malhas fechadas. Nestes casos torna-se impossível um correcto seguimento de contornos e, por isso, o método tende a falhar.

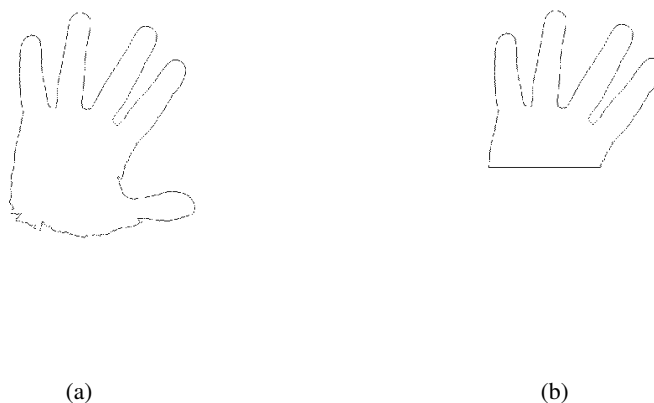


Figura 5.6: a) Imagem original alinhada do contorno; b) Imagem do contorno original cortada ao nível horizontal do CM. Neste caso apenas 4 dedos se encontram na imagem.

5.2.2 Esqueleto da Mão

Um outro método testado para a extracção de pontos característicos foi o recurso ao esqueleto da mão. Neste método é utilizada a máscara binária alinhada da mão e o método baseia-se em determinar o seu esqueleto. Com isso espera-se que cinco das pontas do esqueleto correspondam às pontas dos dedos. A operação de esqueletização reduz um objecto a uma linha, Figura 5.8. Para tal foi utilizada a função *bwmorph()* disponibilizada pelo MatLab. Contudo, nesta função existem dois métodos que se podem ser utilizados mas que apresentam resultados muito semelhantes, são eles o *skel* e o *thin*. O método *skel* funciona retirando os pixels das bordas do objecto até que este seja representado por uma linha enquanto o método *thin* reduz imediatamente o objecto a um



Figura 5.7: Imagem do contorno da parte superior da mão em coordenadas polares (4 dedos apenas).

linha conectada recorrendo a uma divisão do objecto em dois sub-campos. Ambos os métodos preservam o número de *Euler*. Embora pelo nome seja sugestivo escolher-se o primeiro, na verdade o segundo método apresenta melhores resultados, pois exhibe menor número de segmentos na imagem final do que o método *skel*.

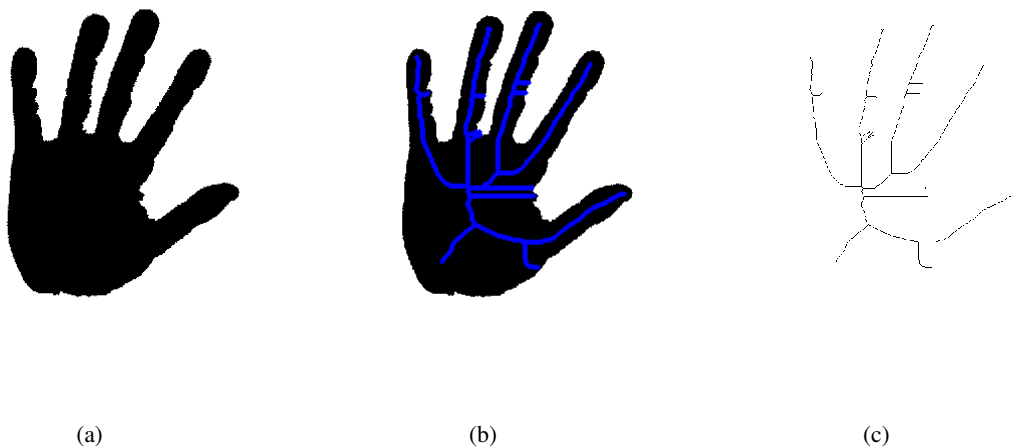


Figura 5.8: a) Máscara binária alinhada da mão; b) Máscara binária com representação do respectivo esqueleto; c) Extração do esqueleto da mão para futuro tratamento.

É, de certa forma, intuitivo notar que alguns dos pontos terminais do esqueleto da mão estão muito próximas às pontas dos dedos (ver Figura 5.8). Esta pequena distância pode ser corrigida, tal como será demonstrado mais à frente. É, no entanto, necessário efectuar primeiro uma operação

morfológica de "poda" para eliminar pequenos segmentos que se encontram no esqueleto e que são indesejáveis. No entanto, esta operação não deve exceder um certo limite de "poda" com o risco de diminuir em demasia os maiores segmentos e, conseqüentemente, tornar muito grande a distância entre a ponta do segmento do esqueleto e a ponta do dedo. Se esta distância se tornar demasiado grande, provocará muitas dificuldades na detecção dos pontos.

Detecção dos pontos correspondentes às pontas dos dedos

Para a detecção dos pontos correspondentes às pontas dos cinco dedos foram utilizadas os pontos terminais de alguns dos segmentos do esqueleto. Tal como foi observado atrás, em todos os dedos existe um segmento que percorre o dedo e termina um pouco antes da sua extremidade. A operação morfológica de "poda" fez com que esta linha fosse única e tornasse assim possível a extracção dos referidos pontos. É necessário, em primeiro lugar, detectar as pontas de cada segmento do esqueleto e identificar aquelas que correspondem a cada dedo. Para tal foi efectuada uma função que detecta todas as pontas do esqueleto e identifica as principais da seguinte forma:

- Os quatro pontos terminais dos segmentos com ordenada menor correspondem aos segmentos que percorrem cada um dos quatro dedos; depois esses pontos são ordenados segundo a abcissa de forma a fazer corresponder cada ponta ao dedo respectivo.
- O ponto terminal do segmento com abcissa maior corresponderá ao segmento que percorre o dedo polegar.

É então utilizado o seguinte método para detecção dos pontos reais:

1. É marcado um ponto **P** com a abcissa do CM mas com ordenada superior à do CM (aproximadamente dez pixels acima do início da mão), $P = [X_{CM}; Y]$;
2. São traçadas sobre a máscara binária cinco linhas que têm início no ponto P e que passam pelo ponto correspondente à ponta do segmento do esqueleto de cada dedo;
3. Cada uma das linhas é percorrida ponto-a-ponto e o primeiro ponto, de cada linha, a ser encontrado não pertencente à mão (1º pixel branco) é marcado como o ponto correspondente à ponta do respectivo dedo;

Tal pode ser observado na Figura 5.9.

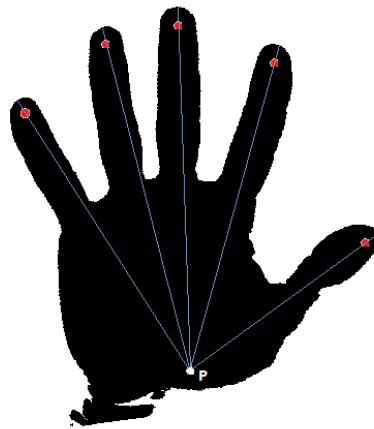


Figura 5.9: Método para detecção dos pontos correspondentes às pontas dos dedos a partir das pontas do esqueleto (pontos marcados a vermelho correspondem às pontas do esqueleto identificadas antes).

Detecção dos pontos correspondentes aos vales entre os dedos

Método que utiliza o esqueleto

Inicialmente pensou-se em utilizar o método baseado no esqueleto para a detecção dos vales. Isto é, inverter a imagem da máscara binária e determinar o esqueleto correspondente. Esperava-se que, de uma forma semelhante à testada para as pontas dos dedos, as pontas do esqueleto correspondessem aos vales. Cedo se verificou que tal aproximação não poderia ser efectuada dada a elevada quantidade de pequenos segmentos presentes no esqueleto. Tal facto pode ser constatado na Figura 5.10. Poderia, tal como anteriormente, ser aplicada uma operação de "poda" para eliminar os pequenos segmentos do esqueleto mas seriam drasticamente diminuídos os segmentos do esqueleto cuja ponta correspondesse ao vale. Seria então extremamente complicado fazer corresponder a ponta do esqueleto ao vale, principalmente devido ao pouco espaço que normalmente existe entre dedos, facto que não se observa nas pontas dos dedos.

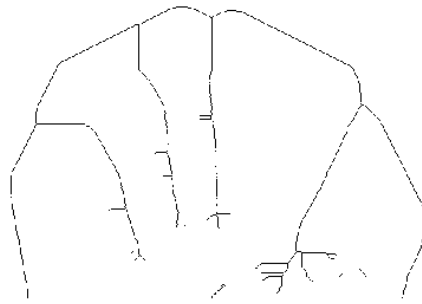


Figura 5.10: Esqueleto da máscara binária da mão invertido para detecção dos pontos correspondentes aos vales entre os dedos.

Método de varrimento horizontal entre linhas que definem os dedos

Dados os resultados negativos obtidos no primeiro método (método do esqueleto), foi implementado um algoritmo para detecção dos pontos correspondentes aos vales entre os quatro primeiros dedos:

1. É traçada uma linha horizontal com ordenada do CM (Y_{CM}), com início e fim no contorno da mão;
2. Esta linha (que tem como comprimento a largura da mão naquela ordenada) é dividida em quatro partes iguais, sendo identificados os três pontos $R_{1,3}$ onde essa divisão é efectuada;
3. Para cada um dos pontos identificados anteriormente são traçadas duas linhas com início nesse ponto e com fim nas duas pontas dos dedos correspondentes;
4. É feito um varrimento horizontal entre essas duas linhas. O primeiro pixel branco detectado corresponderá ao vale entre esses dois dedos;

Para detecção do vale entre o dedo indicador e o polegar foi utilizado um método semelhante mas cujas linhas têm início no ponto \mathbf{P} , atrás referido, e não na linha do CM.

Na Figura 5.11 está ilustrada a forma de detecção dos vales.

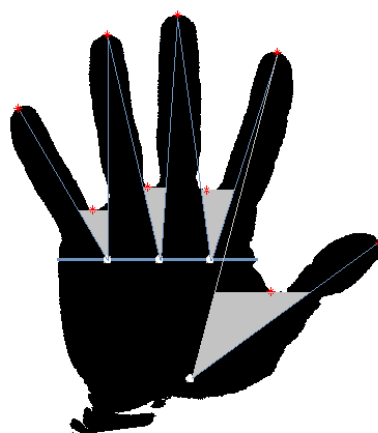


Figura 5.11: Método para detecção dos pontos correspondentes aos vales entre os dedos a partir de varrimentos entre linhas que percorrem os dedos.

Método do seguimento do contorno entre dedos

Apesar dos bons resultados obtidos no método anterior, foi implementado e testado um outro método no âmbito de ultrapassar casos em que a detecção era grosseira. Isto é, casos em que os vales apresentavam muito ruído e o ponto detectado como vale pelo método anterior nem sempre correspondia ao local mais exacto. Neste sentido, um novo método foi desenvolvido em que, dados os pontos correspondentes às pontas dos dedos, é efectuado um seguimento de contornos para detecção dos valores de curvatura em cada trecho de contorno analisado. O funcionamento deste método é descrito de seguida:

1. Identificação dos cinco pontos correspondentes às cinco pontas dos dedos;
2. Efectuar um seguimento do contorno entre dois pontos correspondentes às pontas dos dedos;
3. Aplicar um filtro de suavização para eliminar ruído que possa influenciar a correcta detecção dos pontos;
4. Analisar a curvatura desse trecho de contorno e identificado o ponto com maior valor, recorrendo à análise da primeira derivada;
5. Identificar o pixel do contorno não suavizado que se encontra mais próximo do ponto marcado.

O funcionamento deste método apresenta-se muito simples e oferece excelentes resultados na medida em que detecta sempre os pontos correspondentes aos vales e não está tão sujeito a ruído como os restantes métodos. Contudo, e tal como acontece nos outros métodos, o sucesso deste depende, na totalidade, do sucesso na detecção dos pontos correspondentes às pontas dos dedos. Caso estes falhem, a detecção dos vales irá também ela falhar. Na Figura 5.12 está ilustrado o funcionamento deste método.

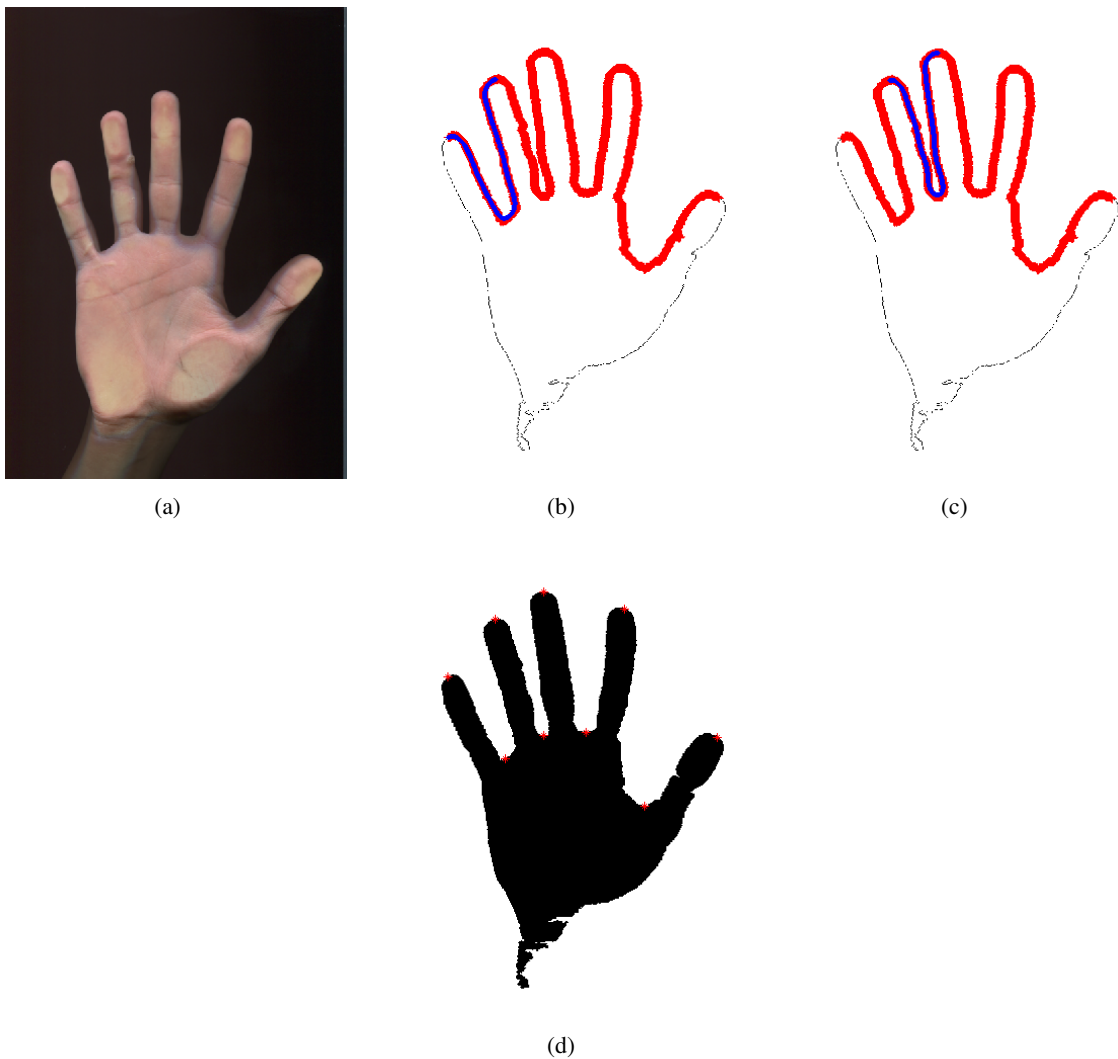


Figura 5.12: a) Imagem original RGB; b) 1º passo do método: seguimento do contorno entre ponta do dedo mínimo e do dedo anelar (contorno original a vermelho, contorno suavizado a azul); c) 2º passo do método: seguimento do contorno entre as pontas dos dedos anelar e médio (contorno original a vermelho, contorno suavizado a azul); d) Resultado final do método.

Resultados

O método baseado no esqueleto, embora estável no que respeita aos grandes segmentos, apresenta muita instabilidade nos pequenos segmentos. Isto é, a operação morfológica de "poda" não pode ser excessivamente extensa pois pode reduzir demasiado os segmentos grandes, mas não existe uma medida que elimine todos os segmentos pequenos e indesejáveis. As principais causas de falha no método que utiliza o esqueleto são:

- Pequenos segmentos de esqueleto que permanecem mesmo após a poda e cuja ponta está numa ordenada menor do que alguma das pontas verdadeiras, (Figura 5.13(a));
- Dedos demasiado juntos que por vezes provocam a junção de dois segmentos do esqueleto quase no final dos mesmos, eliminando assim uma ponta essencial do esqueleto, (Figura 5.13(b));

Relativamente à detecção dos pontos correspondentes aos vales, o segundo método aqui proposto apresentou bons resultados. Embora, em casos com elevado ruído, o ponto marcado esteja ligeiramente afastado do ponto mais realista, este método identifica de uma forma estável todos os vales. No entanto, para ultrapassar a desvantagem induzida pelo ruído no método anterior, surgiu o método do seguimento do contorno entre dedos, que se mostra igualmente estável, mas que apresenta uma maior imunidade ao ruído, oferecendo por isso bons resultados.

A grande desvantagem destes dois últimos métodos reside no facto de que dependem, na totalidade, do sucesso na detecção dos pontos correspondentes às pontas dos dedos. Isto é, caso exista uma correcta identificação das cinco pontas dos dedos, os vales serão correctamente identificados com elevada probabilidade de sucesso. Por outro lado, caso as pontas dos dedos não sejam identificadas correctamente, a detecção dos vales irá falhar.

5.2.3 Polígono Convexo da Mão

O último método testado para a extracção dos nove pontos característicos recorreu ao polígono convexo da mão. Este método foi implementado recorrendo à função disponível no MatLab, *convexhull()*, que retorna a área convexa de um objecto e mostrou ter algum potencial no início. O polígono resultante é compreendido por uma linha que contém como vértices os pontos correspondentes às pontas dos dedos. Pensou-se em determinar os pontos correspondentes a esses vértices e fazê-los corresponder às pontas dos dedos. Deste modo, permitia apenas a detecção dos pontos correspondentes às pontas dos dedos, e não dos vales que poderiam, no entanto, ser detectados a partir de outro processo. Na Figura 5.14 pode ser observada a aplicação deste método.

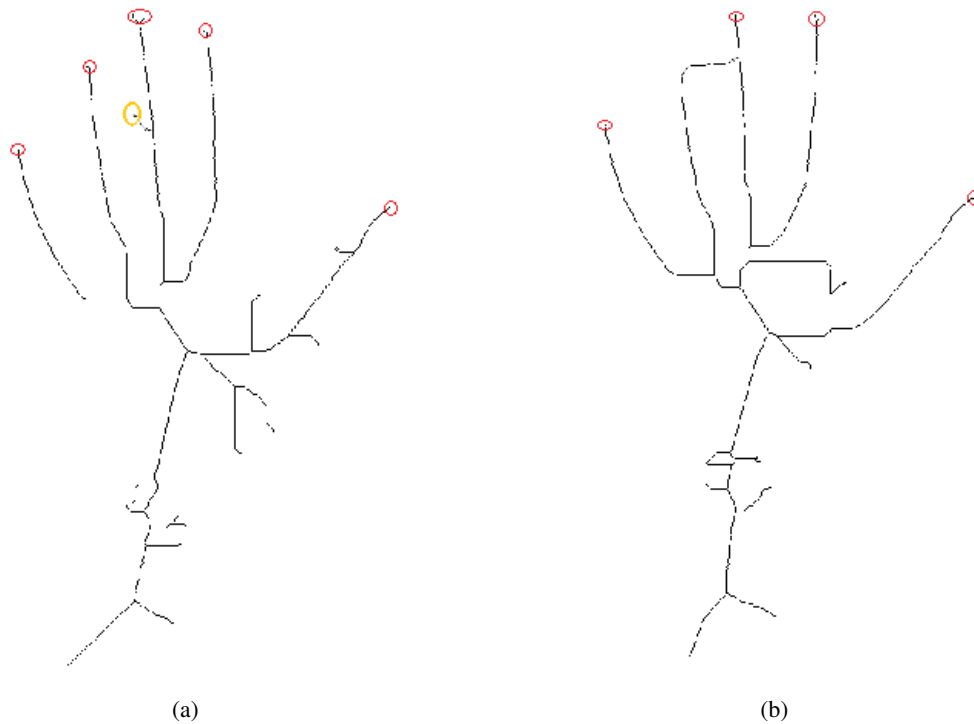


Figura 5.13: a) Esqueleto apresentando pequenos segmentos que permaneceram após a operação de poda e que influenciam a detecção de pontos (pontas marcadas a vermelho correspondem a pontos verdadeiros, ponta marcada a laranja corresponde a ponto marcado erradamente); b) Esqueleto de uma mão em que os dedos se encontram muito juntos, apresentando a junção de dois segmentos em que uma ponta verdadeira é eliminada.

Resultados

Depois de testado este método em todas as imagens da BD, verificou-se que em todas existiam vértices bastantes próximos das pontas dos dedos, podendo por isso ser identificados como sendo pontos característicos. Foi, contudo, observado que para algumas imagens, principalmente aquelas em que a mão se encontra com os dedos fechados, os pontos determinados se encontravam a uma distância considerável dos pontos reais da ponta do dedo. Alguns exemplos disso estão representados na Figura 5.15.

O facto de os pontos identificados como pontas dos dedos estarem a uma distância relativamente elevada do ponto verdadeiro faz com que, na fase de extracção de características, o valor destas esteja influenciado pela detecção errada da localização dos pontos. Então, as medidas não serão correctas e podem até variar dentro do universo de imagens da mesma pessoas devido às

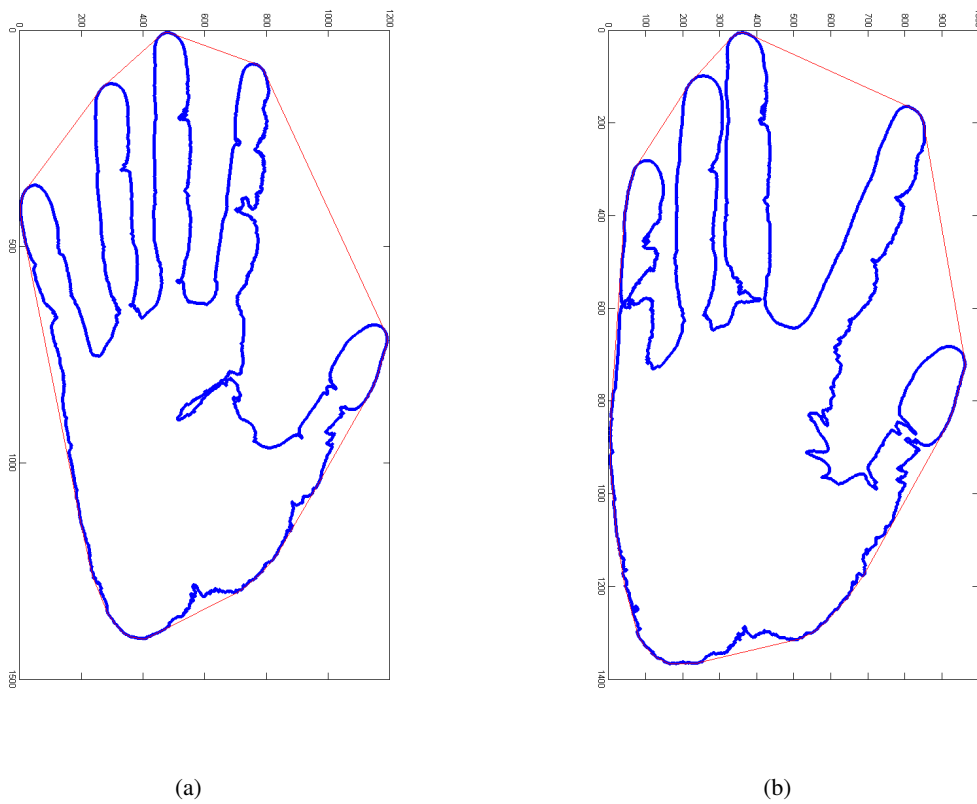


Figura 5.14: a) e b) Área convexa que envolve a mão (a vermelho), resultado da aplicação de *convexhull()* e o respectivo contorno (a azul).

diferentes posições da mão que a mesma pessoa pode tomar. Tendo isto em conta, decidiu-se que este método não verifica as condições necessárias para ser utilizado, pois embora existisse uma elevada taxa de sucesso na detecção dos pontos, a redundante localização dos mesmo iria afectar bastante o sucesso das fases subsequentes do sistema de reconhecimento.

5.3 Conclusão

Após a implementação e o teste de todos os métodos atrás descritos foi necessário fazer uma avaliação do desempenho de cada um deles de modo a seleccionar qual o método mais adequado para a detecção dos nove pontos característicos. Aquele que se mostrou mais eficaz na detecção de todos os pontos, e também o que oferece melhores resultados, foi o método que utiliza o esqueleto da mão para identificação das cinco pontas dos dedos e, para a marcação dos quatro vales, o

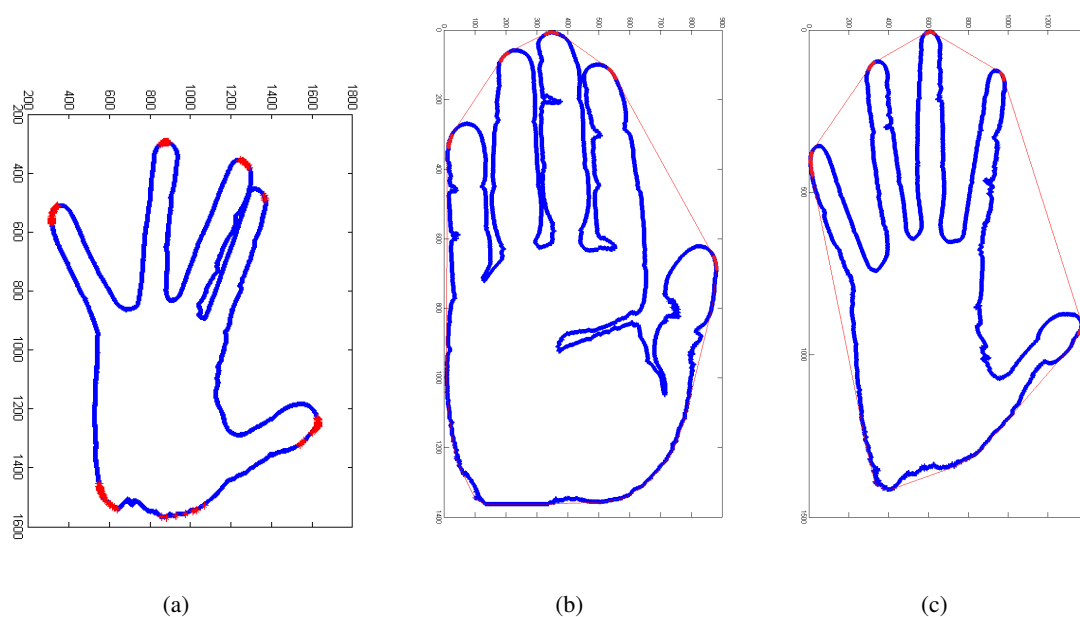


Figura 5.15: a) Marcação dos pontos correspondentes aos vértices da área convexa; b) e c) Identificação dos pontos dos vértices da área convexa que coincidem com as pontas dos dedos.

método do seguimento do contorno entre cada dois dedos. Estes métodos foram aqueles que produziram melhores resultados, tanto quantitativamente (número de pontos detectados) como qualitativamente (distância ao ponto real), principalmente devido à sua imunidade ao ruído. Os resultados deste método podem ser analisados na Tabela 5.1.

Tabela 5.1: Taxas de eficácia na detecção dos pontos característicos da mão.

	<i>BD completa(10 imgs por ind.)</i>	<i>BD reduzida(5 imgs por ind.)</i>	
Número de pontos a identificar	<i>9 pontos</i>	<i>9 pontos</i>	<i>8 pontos</i>
Acerto			
Em pontos	87%	93.4%	98.1%
Em imagens	59%	89.5%	97%

Como pode ser facilmente constatado os resultados não são tão optimistas como esperado. Este facto deve-se, essencialmente, aos casos de insucesso ocorridos na fase de pré-processamento. Em todos os casos em que as imagens apresentavam a mão com os dedos demasiado juntos, a obtenção da máscara binária falhou, consequentemente a extracção do contorno falhou e a detecção dos pontos característicos também falhou. Dados estes casos de insucesso, que provocam uma queda

bastante acentuada na taxa de eficácia de detecção dos pontos e, conseqüentemente, na fase seguinte de extracção de características, chegou-se à conclusão de que a BD não reunia as condições necessárias para que se avançasse para a fase do reconhecimento.

Para o uso pleno da BD seria necessário investir mais tempo na fase de pré-processamento, mais especificamente, no tratamento das imagens que apresentam tais condições. Tal investimento não foi possível devido ao tempo limitado mas ficará registado como trabalho futuro.

Dada esta conclusão, decidiu-se então reduzir a BD retirando as imagens que falhavam na fase de pré-processamento, nomeadamente, aquelas que apresentam a mão com os dedos demasiado juntos. Para que esta nova versão da BD ficasse equilibrada e com o mesmo número de imagens por pessoa, foram retiradas cinco imagens de cada individuo, restando assim outras cinco para as fases seguintes. Apesar de limitada, esta BD apenas apresenta restrição nas imagens com a mão fechada. Todas as restantes em que existe toda a liberdade de rotação e configuração dos dedos foram mantidas, tal como pode ser confirmado na Figura 5.16. Contudo, algumas das imagens que apresentam os dedos juntos, mas que se verificou terem sucesso na fase de pré-processamento, foram também mantidas.



Figura 5.16: Exemplos de imagens que permaneceram na BD depois de reduzida e retiradas aquelas em que a mão se encontrava com os dedos juntos.

Finalmente, analisou-se quantos pontos seriam necessários para posterior extracção de características. Foi observado que nenhum dos trabalhos que definem o estado da arte utilizou as medidas do dedo polegar devido à enorme incerteza da localização correcta do vale entre o dedo indicador e o polegar, que afecta directamente todas as medidas efectuadas no dedo polegar. Logo, este dedo

nunca foi considerado para extracção de características pois prejudicaria a fase de classificação. Embora a sua detecção seja feita aqui, na verdade se for removido o ponto correspondente ao vale, a taxa de eficácia sobe consideravelmente, tal como pode ser verificado também na Tabela [5.1](#).

Capítulo 6

Definição e Extração das Características para o Reconhecimento

6.1 Introdução

Aproximando-se a fase final do sistema de reconhecimento, resta-nos apresentar apenas a extração de características. As características extraídas nesta fase serão as utilizadas na fase de identificação e por isso devem permitir identificar cada pessoa o melhor possível. Também por isso devem ser extraídas da forma mais correcta possível de modo a evitar erros na medição. Segundo vários autores, [Wong e Shi \(2002\)](#); [Tsapatsoulis e Pattichis \(2009\)](#); [Ross et al. \(1999\)](#); [Montalvao et al. \(2010\)](#), entre outros, as características mais relevantes e normalmente utilizadas nestes sistemas são:

- Comprimento dos quatro dedos (dedo mínimo, anelar, médio e indicador);
- Largura de cada dedo (em uma ou mais partes do dedo);
- Forma da região da ponta do dedo;
- Área da mão;
- Perímetro da mão;
- Razão entre maior e menor eixo da elipse que envolve a palma;
- Razão entre área total e área convexa;
- Excentricidade (distância entre foco e eixo principal).

Neste capítulo será apresentada a metodologia para a características necessárias à fase seguinte e será também apresentado um estudo dos atributos extraídos. Este estudo fornecerá alguma informação útil sobre a importância relativa de cada atributo; isto é, quais os que contêm mais informação e quais aqueles que afectam o desempenho do sistema. Desta forma, e com essa informação, será possível saber quais os atributos essenciais à fase de identificação e quais aqueles que podem/devem ser eliminados.

6.2 Métodos e Resultados

6.2.1 Extracção da Área, Perímetro e Excentricidade

A extracção destas primeiras características é, de facto, bastante simples e não requer grandes cálculos nem sequer a implementação de novas funções para tal. Na fase de pré-processamento aquando da fase alinhamento, foi implementada uma função que calcula os momentos da imagem. Sabemos portanto, que o momento $m_{0,0}$ usando a imagem da máscara binária da mão corresponde à área da mão.

Quanto ao perímetro decidi extrair-se duas medidas distintas deste. Uma delas é o perímetro completo da mão, mas, como já pode se observou antes, algumas das imagens apresentam uma significativa parte com a presença do braço. Quando tal acontece o perímetro e área extraídos não correspondem ao real. Decidiu-se então efectuar outra medida de perímetro que não seja tão susceptível de erro nestes casos. A segunda medida é o perímetro da parte superior da mão, ou seja, tal como já foi feito anteriormente a imagem do contorno é cortada ao nível do CM, sendo desta forma calculado o perímetro correspondente à zona dos dedos. Existem duas formas muito simples de obter estas duas medidas. Podemos, tal como para a área, aplicar a função que calcula o momento $m_{0,0}$ na imagem do contorno da mão, obtendo assim a área da imagem do contorno que corresponde ao perímetro da mão. Uma outra forma mais simples de calcular o perímetro é aplicar a função existente no MatLab $nnz()$ que, simplesmente, conta o número de pixels pretos na imagem.

A excentricidade envolve a aplicação de uma função, também ela disponível no MatLab, mas de aplicação muito básica. A função é a $regionprops()$ que, quando usada com os métodos $majoraxislength$ e $minoraxislength$ devolve, respectivamente, o comprimento do maior e do menor eixo da elipse que envolve a zona da mão. Obtidos estes parâmetros, torna-se também fácil o cálculo deste atributo que segue a seguinte equação:

$$eec = \sqrt{1 - \frac{\text{minor_axis_length}^2}{\text{major_axis_length}^2}} \quad (6.1)$$

6.2.2 Extração do Comprimento do Dedo

Uma das medidas com mais importância nestes sistemas é o comprimento de cada um dos dedos. Tal como já foi referido, a marcação do ponto correspondente ao vale entre o indicador e o polegar assume um comportamento bastante inconstante, sendo por isso uma considerável fonte de erro. Assim, as medidas de comprimento e largura do dedo polegar não serão extraídas nem consideradas para efeitos de classificação, tal como se verificou em todos os trabalhos revistos na literatura.

A extração do comprimento de cada dedo revela-se muito simples quando conhecidos os pontos característicos da mão. Para efectuar esta medida são necessários três pontos relativos a cada dedo, são eles os pontos correspondentes à ponta do dedo e aos dois vales que o compreendem. Os pontos correspondentes aos vales que compreendem cada dedo serão aqui mencionados como os pontos da base do dedo. Rapidamente se percebe que faltam dois pontos essenciais que não foram identificados antes, os pontos correspondentes à base esquerda do dedo mínimo e à base direita do dedo indicador. Estes dois dedos apenas têm vales de um dos lados. Será então necessário proceder à detecção desses dois pontos.

Inicialmente pensou-se em fazer uma pesquisa na ordenada do outro ponto base até encontrar um ponto pertencente ao contorno nessa ordenada e marcá-lo como o outro ponto base. No entanto, esta seria uma forma bastante grosseira de o fazer devido ao elevado ruído que o contorno apresenta. Foi então implementado o seguinte método para a detecção desses dois pontos:

1. É, inicialmente, aplicado o algoritmo de seguimento do contorno com início na ponta do dedo e final na ordenada do CM;
2. Durante o seguimento do contorno são guardadas num vector as coordenadas dos pontos percorridos;
3. É calculada a distância Euclideana entre o outro ponto base do dedo e todos os pontos percorridos antes;
4. Aquele que apresentar menor distância é identificado como o ponto que se pretende detectar.

Na Figura 6.1 está ilustrado o método utilizado para a extração destes dois pontos.

Tendo agora identificados todos os pontos necessários, encontram-se reunidas todas as condições para a extração do comprimento de cada dedo. O algoritmo desenvolvido é composto pelos seguintes passos:

1. Para cada dedo i é traçada uma linha entre os dois pontos da sua base;

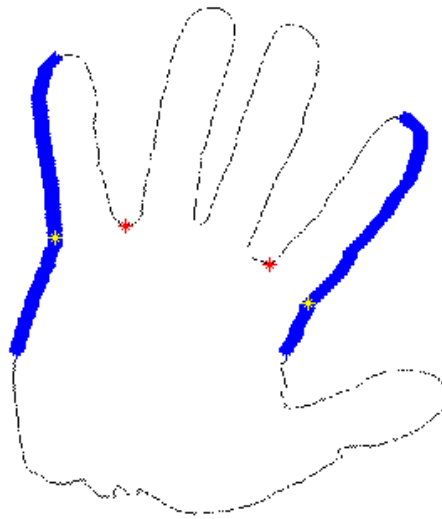


Figura 6.1: Ilustração do método utilizado para detecção dos dois pontos da extremidade da mão correspondentes às bases esquerda do dedo mínimo e base direita do dedo indicador. Pontos a azul representam o trecho do contorno percorrido, pontos marcados a vermelho representam o ponto existente para o qual é calculada a distância aos pontos percorridos e os pontos marcados a amarelo representam os novos pontos identificados.

2. É calculado o ponto médio dessa linha, m_i ;
3. O comprimento de um dedo (distância L_i) corresponde à distância Euclideana entre o ponto médio m_i e a ponta do dedo correspondente.

A Figura 6.2 ilustra o método acima descrito para a extracção do comprimento dos dedos.

6.2.3 Extracção da Largura do Dedo

Os sistemas propostos por Wong e Shi (2002), Ross et al. (1999) e Adan et al. (2008) apresentam duas medidas em zonas diferentes para a largura do dedo, enquanto Faundez-Zanuy e Merida (2005) apresentam um sistema em que é usada apenas uma medida de largura do dedo.

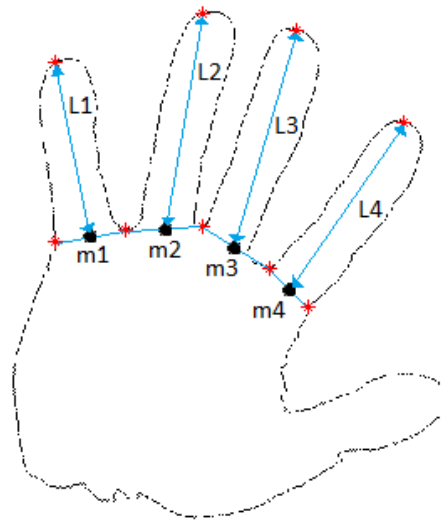


Figura 6.2: Ilustração do método utilizado para extração dos comprimentos $L_{1,4}$.

Foi então necessário decidir sobre a quantidade de medidas de largura a extrair e estipulou-se que se utilizariam duas medidas de largura de cada dedo ($W1_i, W2_i$). Estas medidas correspondem à largura do dedo a $(1/3)$ e $(2/3)$ da distância L_i . Obtidas as medidas do comprimento de cada dedo e os três pontos que os definem (uma ponta e os pontos de base), a extração das larguras é feita de uma forma simples. A medição das larguras é feita seguindo um varrimento horizontal depois do respectivo dedo ser alinhado segundo a direcção vertical. Este método apresenta um melhor desempenho e uma maior simplicidade de implementação relativamente ao método baseado em linhas traçadas segundo o ângulo de inclinação de cada dedo. A metodologia utilizada é descrita de seguida:

1. É extraído da imagem do contorno, o trecho do mesmo que se encontra entre a base esquerda do dedo a tratar e a sua base direita (isolando assim cada dedo do resto do contorno da mão);
2. De seguida, é aplicado o método dos momentos à imagem que contém apenas o contorno do dedo, determinando o seu ângulo de orientação;
3. Obtido o ângulo θ procedeu-se ao alinhamento do contorno do dedo;
4. Agora com o dedo alinhado, são efectuados dois varrimentos horizontais às alturas $(1/3)L_i$ e $(2/3)L_i$.

A Figura 6.3 apresenta o método utilizado para extracção das medidas de largura de cada dedo.

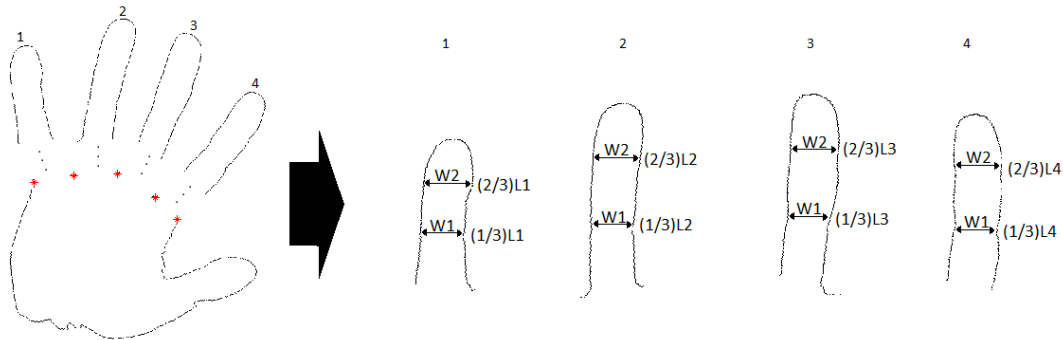


Figura 6.3: Ilustração do método utilizado para determinação das larguras de cada dedo.

6.2.4 Extração da Região da Ponta do Dedo

Aqui decidi não se utilizar uma medida fixa para esta região mas sim, determinar o início da ponta do dedo com base no comprimento do mesmo. Assim sendo, daqui resultou que o mais indicado, e que ao mesmo tempo acarreta mais informação sobre a pessoa, seria definir a região da ponta do dedo como sendo os últimos 20% do comprimento do dedo.

A extração da região da ponta do dedo é feita de uma forma em tudo semelhante ao passo anterior, adoptado para a extração das larguras do dedo. É utilizado o contorno de cada dedo alinhado e a partir daí é extraída a área correspondente à ponta do dedo. Este método de extração da ponta do dedo está demonstrado na Figura 6.4.

Alguns autores, como [Wong e Shi \(2002\)](#) utilizam esta região para fazer *template matching*. Apesar de neste sistema poder ser utilizado o mesmo método, esse não foi o caminho escolhido. Preferiu-se utilizar o método de *feature matching* e, por isso, é necessário extrair alguma informação numérica sobre esta região, nomeadamente a sua área.

6.3 Estudo dos Atributos

Atributos extraídos

Tendo sido implementados todos os métodos necessários, procedeu-se então à extração das características dentro de toda a BD. Um conjunto de atributos obtido para algumas imagens da

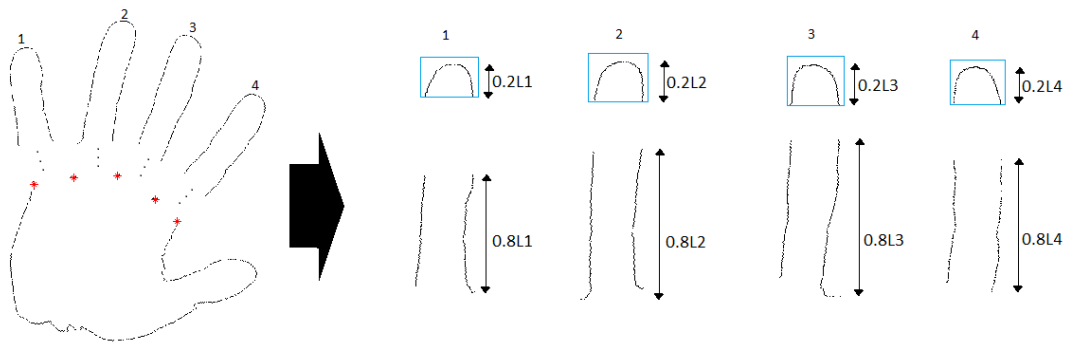


Figura 6.4: Ilustração do método utilizado para delimitação da região da ponta de cada dedo.

BD pode ser observado na Tabela 6.1. De seguida estão também descritos os atributos extraídos e respectiva legenda:

1. Comprimento do dedo mínimo - L1
2. Comprimento do dedo anelar - L2
3. Comprimento do dedo médio - L3
4. Comprimento do dedo indicador - L4
5. Largura a (1/3) do dedo mínimo - W11
6. Largura a (2/3) do dedo mínimo - W12
7. Largura a (1/3) do dedo anelar - W21
8. Largura a (2/3) do dedo anelar - W22
9. Largura a (1/3) do dedo médio - W31
10. Largura a (2/3) do dedo médio - W32
11. Largura a (1/3) do dedo indicador - W41
12. Largura a (2/3) do dedo indicador - W42
13. Área da região da ponta do dedo mínimo - TIP1
14. Área da região da ponta do dedo anelar - TIP2
15. Área da região da ponta do dedo médio - TIP3
16. Área da região da ponta do dedo indicador - TIP4
17. Área da mão - HA
18. Perímetro da mão - HP
19. Perímetro da parte superior da mão (zona dos dedos) - HCP
20. Área/Perímetro - AP
21. Excentricidade - ECC

Tabela 6.1: Excerto do conjunto de atributos extraídos para a fase de reconhecimento (as cinco imagens pertencem à mesma pessoa).

Atributo \ Imagem	img001	img002	img003	img004	img005
L1	121	119	118	118	119
L2	164	166	166	165	166
L3	183	183	184	183	183
L4	166	166	166	166	166
W11	31	32	33	32	32
W12	26	28	29	28	28
W21	27	31	32	31	31
W22	29	31	30	31	31
W31	31	34	34	33	34
W32	34	34	34	35	34
W41	32	34	33	37	34
W42	34	35	36	35	35
TIP1	57	55	56	54	55
TIP2	77	80	78	79	80
TIP3	84	86	86	85	86
TIP4	81	79	78	78	79
HA	72352	67695	71851	72954	67695
HP	2447	2359	2478	2468	2359
HCP	1751	1603	1690	1694	1603
AP	29.60	28.70	29.00	29.60	28.70
ECC	0.89	0.81	0.87	0.88	0.81

Preparação do conjunto de dados

A preparação do conjunto de dados refere-se à sua análise e à dos seus atributos para detectar algumas discontinuidades que influenciem de forma errada as fases seguintes. Assim sendo, foi realizada uma análise estatística dos atributos que compõem o conjunto de dados tendo sido verificados alguns valores de cada grupo de atributos, como a média, o desvio padrão e os valores máximo e mínimo de cada um deles. Constatou-se então que, provavelmente devido a alguma falha na detecção dos pontos característicos, existiam alguns *outliers*, correspondendo um *outlier* a uma ocorrência de um atributo com valores que se encontram muito distantes da média correspondente, tendo-se procedido à eliminação de tais observações. Após esta análise e preparação do conjunto de dados, estão reunidas as condições necessárias para prosseguir para a fase seguinte. Na Tabela 6.2 podem ser observadas as características estatísticas de todos os atributos de toda a BD após a preparação do conjunto de dados.

Tabela 6.2: Características estatísticas dos atributos extraídos.

Atributo	Média	Desvio Padrão	Máximo	Mínimo
L1	131.86	13.44	164	94.90
L2	173.33	13.47	199	139
L3	185	13.74	215	145
L4	165.11	13.63	193	109
W11	37.23	4.56	51	27
W12	37.46	4.94	47	24
W21	38.08	4.36	50	27
W22	40.94	4.72	52	29
W31	39.59	4.26	50	26
W32	42.75	5.31	55	31
W41	39.23	4.39	51	30
W42	41.67	5.31	56	30
TIP1	64.10	7.36	81	46
TIP2	83.63	6.79	99	60
TIP3	89.02	7.21	104	68
TIP4	79.04	6.42	94	52
HA	79896.75	12730.18	114195	48001
HP	2291.73	256.73	2934	1678
HCP	1504.72	167.47	1859	1082
AP	34.75	2.87	41.80	27.80
ECC	0.72	0.09	0.93	0.51

Estudo da importância relativa de cada atributo

Extraídos e tratados todos os atributos, interessa agora perceber qual a importância relativa de cada um. Isto é, perceber o contributo de cada um deles para o reconhecimento e a sua consistência. Pretende-se, deste modo, seleccionar os atributos que melhor discriminam o resultado da classificação e retirar aqueles que não apresentam esta característica. Para tal, foi efectuado um estudo desta importância recorrendo à análise da correlação existente entre os atributos e entre os atributos e as classes. A medida utilizada foi a correlação de *Pearson* e os resultados para a correlação *atributos vs classes* podem ser vistos na Tabela 6.3.

Analisando os resultados obtidos da medida de correlação de *Pearson*, podem ser observadas as dependências entre cada um deles (por motivos de espaço serão aqui apenas representados alguns desses valores). Pode ser facilmente constatado, na Tabela 6.4, que os comprimentos dos quatro dedos (L_i) estão bem correlacionados entre si assim como com a largura de cada um, Tabela 6.5. Esta conclusão é, de certa forma, lógica pois, para a mesma pessoa, o comprimento e largura do dedo mínimo está relacionado com o comprimento e largura dos restantes dedos, pois dependem todos do tamanho, forma, área e perímetro da mão. Observam-se também altos índices de correlação entre o comprimento e área da ponta do dedo o que, de certa forma, toma algum significado dado serem medidas dependentes entre as mesmas, tais resultados podem ser observados na Tabela 6.6.

Tabela 6.3: Resultados da medida de correlação de *Pearson* para: Atributos vs Classes. Atributos a verde: bem correlacionados; Atributos a vermelho: correlacionados negativamente; Atributo a cinzento: bem correlacionado mas devido à sua elevada escala infere instabilidade na classificação.

Atributo	Correlação de Pearson
L1	0.35
L2	0.11
L3	0.04
L4	0.07
W11	0.34
W12	0.24
W21	0.31
W22	0.15
W31	0.17
W32	0.01
W41	0.12
W42	0.09
TIP1	0.37
TIP2	0.16
TIP3	0.10
TIP4	0.07
HA	0.15
HP	-0.01
HCP	-0.03
AP	0.28
ECC	-0.28

Tabela 6.4: Resultados da medida de correlação de *Pearson* entre os quatro dedos.

Atributo	L1	L2	L3	L4
L1	1	0,80	0,68	0,60
L2	0,80	1	0,93	0,82
L3	0,68	0,93	1	0,91
L4	0,60	0,82	0,91	1

Analisando agora a Tabela 6.3 pode ser observada a importância relativa de cada atributo face às classes. Como resultado desta análise podem ser descartados alguns atributos e pode ser estabelecida uma ordem de importância. Tal pode ser observado na Tabela 6.7. A ordem de importância é facilmente obtida a partir da ordenação dos valores de correlação de cada atributo. Assim, ao

Tabela 6.5: Resultados da medida de correlação de *Pearson* entre comprimento e largura de cada dedo.

Atributo	W11	W12	W21	W22	W31	W32	W41	W42
L1	0,65	0,57	0,66	0,39	0,53	0,34	0,46	0,33
L2	0,67	0,56	0,66	0,48	0,61	0,48	0,59	0,42
L3	0,63	0,50	0,61	0,44	0,60	0,52	0,58	0,42
L4	0,55	0,38	0,54	0,41	0,52	0,43	0,50	0,38

Tabela 6.6: Resultados da medida de correlação de *Pearson* entre comprimento e área de cada dedo.

Atributo	TIP1	TIP2	TIP3	TIP4
L1	0,95	0,80	0,69	0,58
L2	0,77	0,94	0,89	0,80
L3	0,67	0,90	0,94	0,89
L4	0,60	0,79	0,86	0,94

analisarmos os referidos valores, facilmente se constata que a excentricidade e os dois valores do perímetro não contribuem para uma boa classificação, pois apresentam valores inversamente correlacionados. Embora a área da mão apresente um valor de correlação considerável, foi observado ao longo das experiências, que este atributo não contribui para melhorias na identificação. Isto deve-se ao facto de este atributo tomar valores muito elevados comparativamente aos restantes atributos. No final, este atributo irá, de certa forma, minimizar a informação dos restantes. Foi então decidido retirar os quatro atributos referidos para meios de classificação.

Uma outra forma de determinar a importância relativa de cada atributo é recorrendo à função disponibilizada pelo MatLab, *varimportance()*. Esta função estima a importância de cada atributo durante a construção de uma árvore de decisão para a respectiva classificação. Esta função calcula estimativas da relevância de cada atributo somando as alterações no risco, devido à divisão em cada atributo. Em cada nó da árvore de decisão, o risco é estimado como sendo a impureza ou o erro do nó. Esse risco é pesado com base na probabilidade desse nó. A importância associada a esta divisão é então calculada como a diferença entre o risco do nó-pai e o risco total para os dois nós-filhos. O resultado da aplicação desta função pode ser verificado também na Tabela 6.7.

Normalização dos dados

Após todo o tratamento efectuado ao conjunto de dados, foi colocada a hipótese da sua normalização. No entanto, considera-se que a utilização dos dados com os seus valores reais seja mais adequada e mais realista, devido às grandes discrepâncias existentes entre as escalas dos diferentes atributos. Pensa-se, por isso, que o conjunto de dados nesta forma contém mais informação.

Tabela 6.7: Lista dos atributos ordenada pela importância dos mesmos, utilizando dois métodos distintos.

<i>Ordem</i>	<i>Ordem standard</i>	<i>Análise de Correlações</i>	<i>varimportance()</i>
1º	L1	TIP1	L2
2º	L2	L1	L4
3º	L3	W11	HA
4º	L4	W21	L3
5º	W11	AP	TIP2
6º	W12	W12	W11
7º	W21	W31	W32
8º	W22	TIP2	L1
9º	W31	W22	AP
10º	W32	W41	HP
11º	W41	L2	W31
12º	W42	TIP3	W41
13º	TIP1	W42	TIP1
14º	TIP2	TIP4	ECC
15º	TIP3	L4	W22
16º	TIP4	L3	-
17º	HA	W32	-
18º	HP	-	-
19º	HCP	-	-
20º	AP	-	-
21º	ECC	-	-

Contudo, para posteriores experiências decidiu construir-se um conjunto de dados semelhante ao existente, mas normalizado. A normalização foi feita de modo a que cada atributo tomasse valores entre zero e um, utilizando a seguinte fórmula:

$$ValorNormalizado = \frac{Valor_{actual} - min}{Max - min} \quad (6.2)$$

em que *Max* e *min* representam os valores máximo e mínimo do atributo em questão.

É também sabido que a normalização deve ser feita de forma a que informação do conjunto de treino não seja transportada para o conjunto de teste, para não influenciar posteriores resultados da classificação. Assim, o conjunto de dados inicial foi dividido em dois (parte de treino e parte de teste). O conjunto de treino foi normalizado segundo a fórmula indicada acima. O conjunto de teste foi normalizado também segundo a fórmula acima mas com os valores de máximo e mínimo de cada atributo utilizados na normalização do conjunto de treino. Desta forma, nenhuma informação do conjunto de treino está representada no conjunto de teste.

6.4 Conclusão

Neste capítulo foram apresentados e descritos todos os métodos utilizados na extração das características úteis para a construção do conjunto de dados. Foram também descritos todos os atributos que constituem este conjunto. Numa segunda fase, já com o conjunto de dados completo procedeu-se à sua preparação eliminando as observações em que se verificavam descontinuidades que poderiam introduzir erro na fase da classificação, através de uma análise estatística de cada atributo. Foi também efectuado um estudo das características de forma a determinar uma ordem de importância dos atributos. Neste estudo foi também possível identificar aqueles atributos que contêm menor relevância para o processo de classificação. Por fim, foi feita uma normalização do conjunto de dados para futuras experiências e a comparação entre conjunto de dados não normalizado e normalizado.

Resultam então deste capítulo dois conjuntos de dados: um normalizado e outro com os valores reais. Resultam também duas ordens de importância dos atributos obtidas a partir de dois métodos diferentes. Na fase seguinte, o reconhecimento, serão testados os conjuntos de dados e as ordens de importância de forma a seleccionar qual o melhor conjunto de dados e a ordem de importância que produz melhores resultados.

Capítulo 7

Identificação e Reconhecimento

7.1 Introdução

Este capítulo apresenta a fase final do sistema desenvolvido, a fase de reconhecimento. O reconhecimento tem como objectivo classificar cada uma das imagens baseado em informações estatísticas extraídas anteriormente. Um sistema completo de reconhecimento como o nosso consiste numa fase de aquisição das imagens, um mecanismo de extracção de características que retira informações numéricas ou simbólicas das observações e um esquema de classificação das observações, que depende das características extraídas. O esquema de classificação é geralmente baseado na análise de um conjunto de atributos que foram anteriormente classificados, o "conjunto de treino". A classificação é, neste caso, supervisionada uma vez que o sistema conhece o conjunto de saída dos casos a tratar, estabelecendo então os resultados através da análise de erro.

Concluídas as primeiras fases do sistema de reconhecimento, correspondentes à aquisição e extracção de características, falta agora definir a fase do reconhecimento/identificação. Nesta fase, serão expostos dois meios distintos para a identificação: um deles em que se recorrerá a algoritmos de aprendizagem automática e outro em que será utilizado um método de correspondência através do cálculo de distâncias. Na realidade, existem vários factores que decidem a utilização de um ou de outro. A grande vantagem da utilização de um método de correspondência por distâncias reside no facto de o sistema não ter a necessidade de ser re-treinado em todas as vezes que seja introduzido um novo utilizador na BD. Na verdade, esta vantagem passa a ser uma desvantagem quando se opta por um método baseado em aprendizagem automática. Contudo, os algoritmos de aprendizagem como métodos de classificação têm também as suas mais-valias. Por exemplo, pode ser implementado num sistema de reconhecimento biométrico um método de classificação baseado em algoritmos de aprendizagem e, cada vez que um utilizador tenta o reconhecimento e este resulte positivo, o sistema é treinado novamente incluindo a imagem adquirida naquele instante como parte do conjunto de treino. Assim, podem ser vencidos factores como o envelhecimento, novas

marcas/cicatrices contraídas pelos utilizadores, entre outros factores. Existe assim uma adaptação do sistema de classificação ao universo em que está inserido.

7.2 Métodos e Resultados

Dado o reduzido número de imagens por indivíduo presente na BD (cinco/quatro/três imagens por pessoa), recorreu-se a duas formas de separação do conjunto de dados (treino e teste). Uma das formas consistiu em seleccionar aleatoriamente quatro vectores de características por indivíduo para treino e um para teste. Este conjunto será, de aqui em diante, chamado de **”conjunto fixo”**. A outra forma de separação baseou-se numa escolha aleatória dos vectores de características e do seu número representativo na BD, ou seja:

- Para classes que contenham cinco vectores de características, podem ser seleccionados, aleatoriamente, dois, três ou quatro para treino e os restantes para teste;
- Para classes que contenham quatro vectores de características, podem ser seleccionados, aleatoriamente, dois ou três para treino e os restantes para teste;
- Para classes que contenham apenas três vectores de características, podem ser seleccionados, aleatoriamente, um ou dois para treino e os restantes para teste;

Este conjunto será, de aqui em diante, chamado de **”conjunto aleatório”**.

7.2.1 *Feature Matching*

O primeiro método aqui exposto será o *Feature Matching*, que corresponde a um tipo de classificação baseada na correspondência através de distâncias. Este método tem um funcionamento bastante simples e intuitivo:

1. Feita a aquisição da imagem a classificar e extraídas as respectivas características, é calculada a distância Euclidiana entre o vector de características desta imagem e os vectores de características correspondentes a todas as classes presentes no conjunto de treino;
2. A classe atribuída será aquela que obtiver a menor distância.

Foi também testado este método usando os valores da correlação entre o vector de características actual e os presentes na BD. Contudo, em todos os testes efectuados, o desempenho obtido foi significativamente inferior ao obtido pela distância Euclidiana e, desta forma, decidiu-se não utilizar esta métrica no *feature-matching*.

Construído o algoritmo que implementa os passos acima descritos, passou-se ao teste sob o conjunto das imagens de treino. Existem vários testes que podem ser efectuados e grande parte

deles foram, aqui, realizados. Foram efectuados testes com ambos os conjuntos de treino e teste (conjunto fixo e aleatório) em diferentes ordens de inserção de atributos.

Dado existir, uma ordem de importância relativa de cada atributo (Secção 6.3), proveniente do capítulo anterior, resolveu efectuar-se algumas séries de 50 experiências de classificação em que são utilizados os conjuntos fixo e aleatório e em que são introduzidos, iterativamente, atributo a atributo. Desta forma pode perceber-se a evolução da taxa de eficácia na classificação, à medida que se iam introduzindo os diferentes atributos no algoritmo de *feature matching*. Espera-se que a taxa de evolução tenha um aspecto parabólico, para que, desta forma, se possa perceber qual o número de atributos a inserir de modo a ser obtida a melhor taxa de acerto. As 50 experiências referidas foram efectuadas para três ordens distintas de inserção de atributos: uma foi a ordem referida (Secção 6.3) e que foi resultado da análise de correlações; outra foi uma ordem aleatória, que era gerada novamente em cada iteração pela função *rand()*; e a última foi a ordem *standard* (Tabela 6.7).

A Tabela 7.1 apresenta os resultados da *Genuine Acceptance Rate* (GAR) para os diferentes testes. Note-se que, no conjunto fixo, a cada iteração são gerados novos conjuntos de treino e teste que, embora tenham um número fixo de vectores para treino e teste, estes não serão sempre os mesmos. A ordem aleatória é também gerada novamente a cada iteração e utiliza os 21 atributos tal como a ordem original, a ordem das correlações utiliza apenas 17 atributos.

Tabela 7.1: Resultados para 50 experiências variando, em todas elas, os conjunto de treino e teste e a ordem de inserção dos atributos - *feature matching*.

Conj. Utilizado	Ordem das correlações		Ordem aleatória		Ordem original	
	Máximo (%)	Média (%)	Máximo (%)	Média (%)	Máximo (%)	Média (%)
Conj. aleatório	86.00	76.20	86.00	75.60	-	-
Conj. fixo	92.50	83.60	95.00	83.80	86.30	76.50

A título de exemplo, a Tabela 7.2 e a Figura 7.2 mostram os resultados para uma das 50 experiências, em que foi utilizado o conjunto aleatório com a ordem de inserção determinada a partir da medida das correlações. Pode ser verificado que a taxa de eficácia tende a assumir um comportamento parabólico pois, após obtido o máximo, esta começa a decrescer. Tal não pode aqui ser claramente evidenciado pois apenas são usados 17 atributos. Caso se utilizassem os 21 atributos seria mais evidente a parábola que a taxa de eficácia descreve, tal como se pode comprovar pela Figura 7.1.

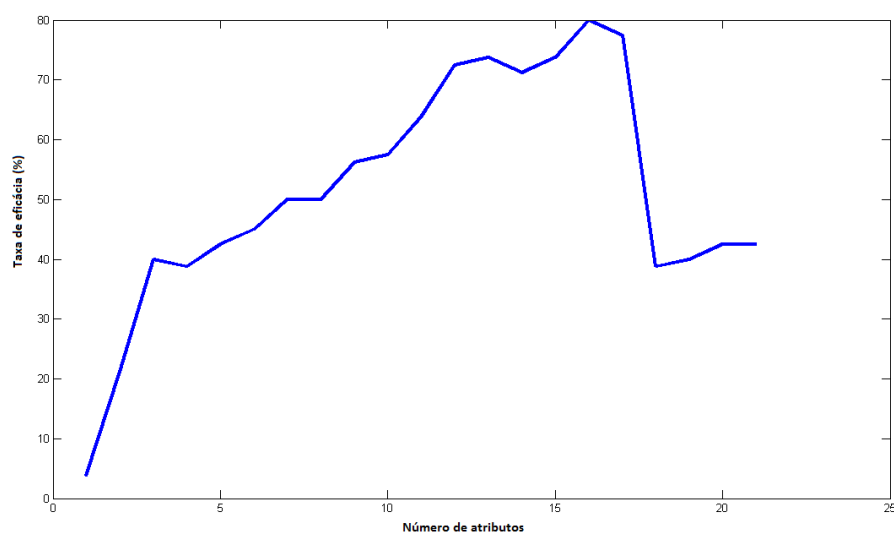


Figura 7.1: Gráfico representativo do comportamento da taxa de eficácia obtida pelo método de *feature matching* introduzindo um atributo de cada vez, pela ordem obtida através das correlações.

Tabela 7.2: Resultados para uma das 50 experiências utilizando conjunto aleatório com ordem de inserção determinada pela análise das correlações - *feature matching*. Para esta experiência teve-se: 110 imagens para treino e 77 imagens para teste.

<i>Iteração</i>	<i>Taxa de eficácia (%)</i>	<i>Atributo inserido</i>
1	7.14	TIP1
2	25.71	L1
3	41.43	W11
4	41.43	W21
5	47.14	AP
6	47.14	W12
7	55.71	W31
8	58.57	TIP2
9	61.43	W22
10	62.86	W41
11	68.57	L2
12	70.00	TIP3
13	71.43	W42
14	75.71	TIP4
15	78.57	L4
16	81.43	L3
17	82.86	W32

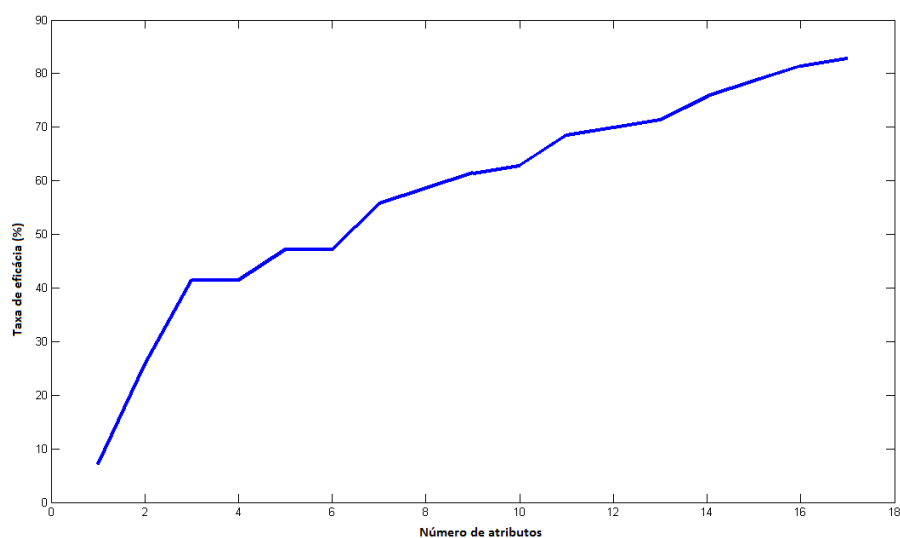


Figura 7.2: *Feature matching* - Gráfico representativo da evolução da taxa de eficácia obtida na experiência presente na Tabela 7.2.

Realizados todos os testes, falta apenas implementar um *threshold* que introduza no sistema uma taxa de rejeição. Isto é, até ao momento qualquer utilizador que tente a identificação neste sistema seria identificado como alguém da BD, mesmo sendo um impostor. É então necessário implementar um método que apenas identifique determinado indivíduo se a distância obtida pelo *feature matching* estiver abaixo de um determinado valor de *threshold*. Foram então efectuados testes de identificação com indivíduos que foram previamente eliminados da BD original e foi feita uma variação do valor de aceitação (*threshold*). Para cada valor de *threshold* testado, foi obtida uma taxa de eficácia GAR diferente. É possível agora apresentar uma taxa de rejeição deste sistema (False Acceptance Rate - FAR) para cada valor de *threshold* tomado. Os valores referidos encontram-se na Tabela 7.3.

Calculados os valores da FAR e GAR para os diferentes valores de *threshold* é agora possível representar um gráfico interessante e bastante comum neste tipo de sistemas. Este gráfico ilustra a variação da FAR e da FRR (False Rejection Rate = 1-GAR) e o ponto onde ambas as curvas se cruzam indica o Equal Error Rate - EER. A partir deste gráfico, pode ser identificado o valor de *threshold* a utilizar no sistema que combina a menor FAR com a maior GAR (Montalvão et al., 2010). A Figura 7.3 ilustra o gráfico referido.

Um aspecto ao nível do funcionamento do sistema quando utilizado o método de *feature matching* para classificação pode ser observado no gráfico da Figura 7.4. A curva ROC aí representada permite um entendimento mais claro e objectivo do funcionamento do sistema. Esta curva

Tabela 7.3: Resultados da GAR e FAR para os diferentes valores de *threshold - feature matching*.

Valor de <i>threshold</i>	GAR (%)	FAR (%)	FRR = (1 - GAR) (%)
1	30	0	70
2	30	0	60
3	40	0	52.50
4	47.50	0	42.50
5	57.50	0	42.50
6	60	0	40
7	60	0	40
8	65	0	35
9	67.5	0	32.50
10	67.50	4.30	32.50
11	72.50	8.70	27.50
12	72.50	8.70	27.50
13	75	26.10	25
14	75	30.40	25
15	77.50	30.40	22.50
16	80	47.80	20
17	80	52.20	20
18	80	56.50	20
19	80	65.20	20
20	82.50	82.60	17.50

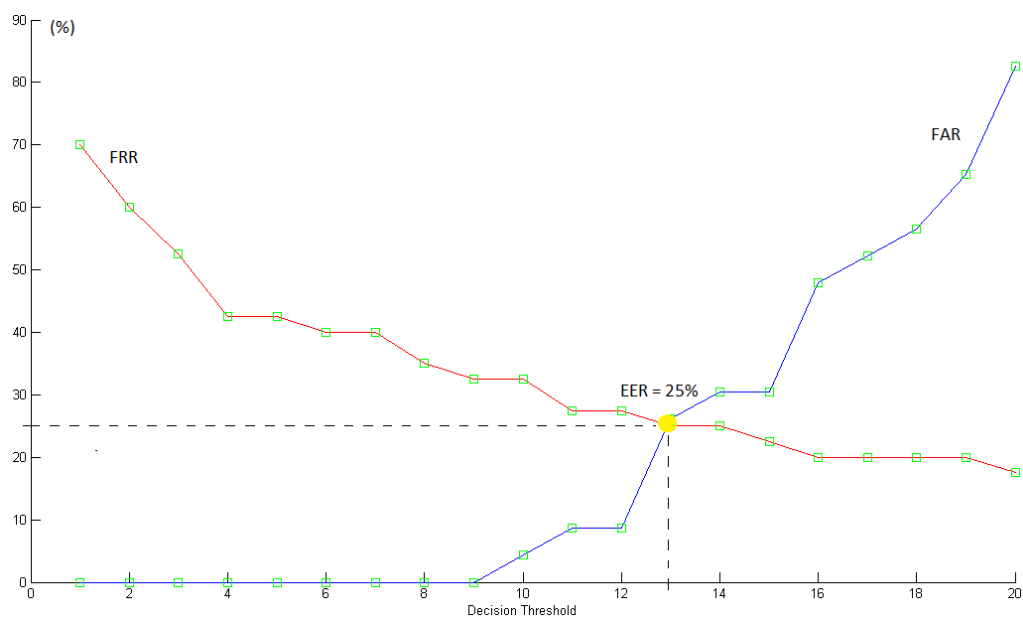


Figura 7.3: FAR vs FRR. Curva a Vermelho - FRR; Curva a azul - FAR; Ponto marcado a amarelo - EER.

apresenta a taxa de eficácia na identificação de pessoas, GAR, em função da taxa de intrusão no sistema, a FAR. O sistema pode funcionar em qualquer ponto desta curva, dependendo da aplicação a que se destina. Se a aplicação final é de elevada segurança então o sistema deve funcionar

próximo do ponto **A**, onde a taxa de intrusão é mais baixa, mesmo que isso implique uma GAR baixa. Se a aplicação final é de baixo nível de segurança, então o sistema deve funcionar próximo do ponto **B**, onde a taxa de eficácia é maior, mesmo a FAR assumindo valores mais elevados. A grande desvantagem do funcionamento em pontos próximos de **A** é que se torna incómodo, para aplicações de pouca segurança, o mesmo utilizador poder ter de tentar a identificação várias vezes, devido à baixa, ou quase nula, FAR.

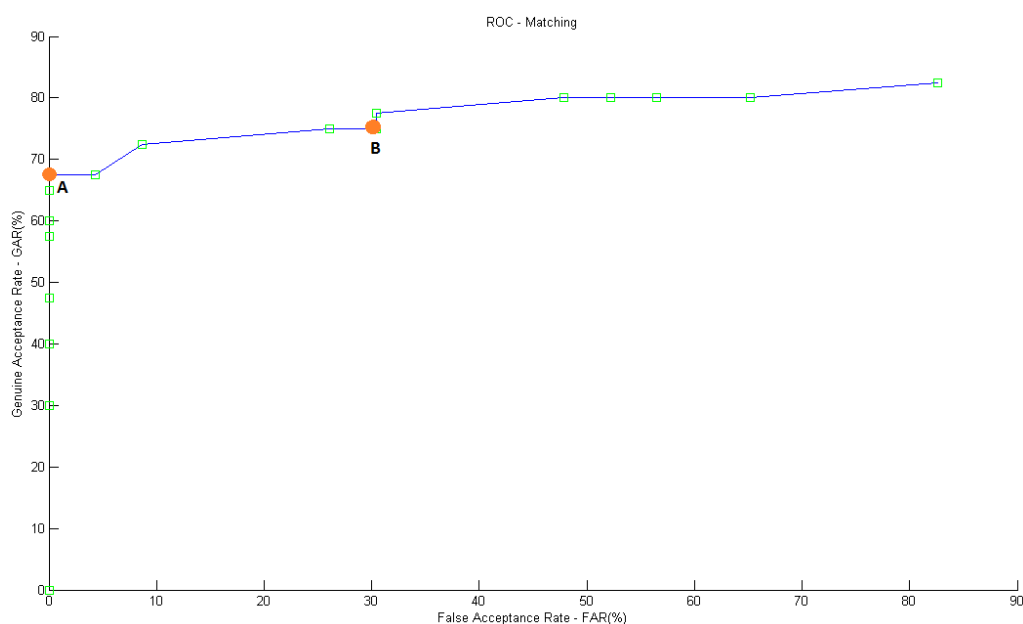


Figura 7.4: Curva de funcionamento do sistema ROC - Receiver Operator Characteristic. Ponto A: funcionamento para aplicações de elevada segurança; Ponto B: funcionamento para aplicações de baixa segurança.

7.2.2 SVM - Support Vector Machine

Os SVMs são um conjunto de métodos de aprendizagem supervisionada que recorrem a uma análise de dados para reconhecimento de padrões. Estes métodos são normalmente usados para aplicações de classificação e regressão. Um SVM padrão analisa um conjunto de dados de treino e prevê, a qual das duas possíveis classes de saída pertence uma determinada amostra. Isto faz do SVM um classificador binário, não probabilístico no qual podem ser utilizados diferentes tipos de funções (*kernel*) tais como linear, Radial Basis Function (RBF), polinomial, entre outros.

Um modelo SVM é então uma representação dos exemplos como pontos num determinado espaço, mapeados de forma a que os exemplos de classes distintas sejam divididos por um intervalo

tão amplo quanto possível. Novos exemplos são então mapeados no mesmo espaço e previstos pertencer a determinada classe com base no lado do intervalo a que pertencer.

Explicado um pouco da teoria deste tipo de classificadores e analisados os resultados que estes oferecem, recorrendo a exemplos que os utilizam, eles tornaram-se bastante atractivos para uso neste sistema. Contudo, como se pode verificar na descrição teórica acima, estes classificadores são binários, ou seja, apenas classificam as várias observações em duas classes (zero ou um, por exemplo). Existem, no entanto, duas grandes abordagens quando se pretende aplicar um classificador SVM em sistemas multi-classe: a abordagem "um contra um" e "um contra todos". Neste caso será aplicada a abordagem "um contra um" que se encontra implementada pelo *libSVM* (Chang e Lin, 2001).

Implementado o classificador SVM, procedeu-se aos vários testes de forma a comprovar os resultados. Os testes aqui realizados foram em tudo semelhantes aos realizados no método *feature matching*. Assim, foram realizadas 50 experiências com duas ordens de entrada de atributos, a ordem resultante da análise de correlações a uma ordem gerada aleatoriamente a cada experiência. Foram também efectuados estes testes nos dois conjuntos de dados, o conjunto fixo e o conjunto aleatório. Tais resultados podem ser comprovados na Tabela 7.4 e os resultados de uma das 50 experiências podem também ser observados na Tabela 7.5 e na Figura 7.5. Do resultado das 50 experiências pode concluir-se que a taxa de eficácia assume um comportamento sempre crescente, à medida que são introduzidos os atributos, atingindo o seu valor máximo na última iteração.

Tabela 7.4: Resultados para 50 experiências variando, em todas elas, os conjunto de treino e teste e a ordem de inserção dos atributos - SVM.

	Ordem das correlações(17att)			Ordem aleatória(21att)	
	<i>Kernel</i>				
	<i>Linear</i>		<i>RBF e Polinomial</i>	<i>Linear</i>	
<i>Conjunto utilizado</i>	<i>Máximo (%)</i>	<i>Média (%)</i>	<i>Máximo (%)</i>	<i>Máximo (%)</i>	<i>Média (%)</i>
Conjunto aleatório	85.30	76.10	63.10	85.40	76.50
Conjunto fixo	84.90	77.80	82.50	88.60	76.60

Tal como no *feature matching* seria agora altura de implementar um método de rejeição e, dessa forma, calcular a respectiva taxa de intrusão no sistema, a FAR. No entanto, embora sendo possível configurar o SVM para devolver saídas probabilísticas, estas assumem intervalos próprios em cada observação, impedindo assim uma implementação eficiente de uma metodologia consistente de *threshold*. Não foi por isso testado o classificador quanto à possibilidade de intrusão por parte de impostores e não é possível a apresentação dos valores da FAR e FRR relativas a este método de classificação.

Tabela 7.5: Resultados para uma das 50 experiências utilizando conjunto fixo com ordem de inserção determinada pela análise das correlações - SVM.

<i>Iteração</i>	<i>Taxa de eficácia (%)</i>	<i>Atributo inserido</i>
1	13.40	TIP1
2	21.00	L1
3	38.80	W11
4	41.80	W21
5	46.30	AP
6	53.70	W12
7	52.20	W31
8	59.70	TIP2
9	68.70	W22
10	64.20	W41
11	65.70	L2
12	76.10	TIP3
13	76.10	W42
14	76.10	TIP4
15	80.60	L4
16	80.60	L3
17	82.10	W32

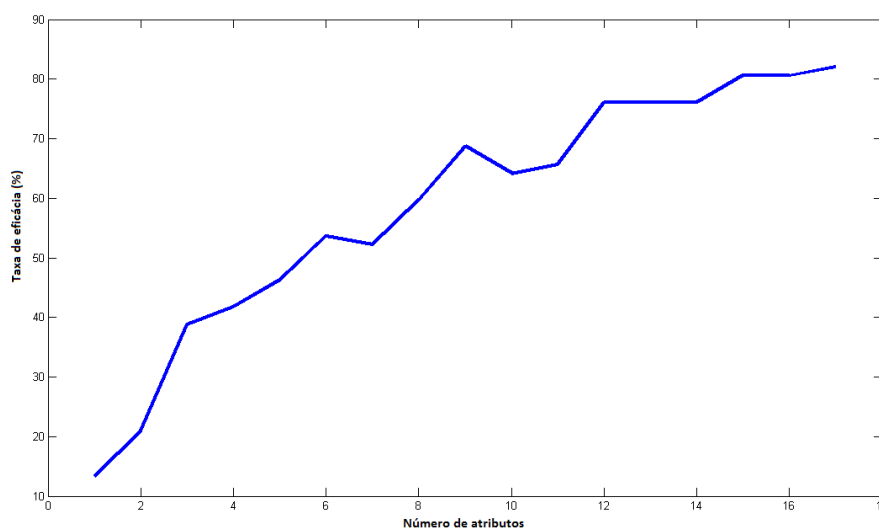


Figura 7.5: SVM - Gráfico representativo da evolução da taxa de eficácia obtida na experiência presente na Tabela 7.5.

7.2.3 Árvores de Decisão

Modelos baseados em árvores de decisão são modelos simples, mas muito usados, que funcionam através da repartição do espaço de entrada em regiões de formato cúbico, cujas arestas são alinhadas com os eixos, e atribuindo um modelo simples (por exemplo, uma constante) a cada região (Bishop, 2006). Uma propriedade chave dos modelos baseados em árvore de decisão é que

estes são facilmente interpretáveis por seres humanos, porque correspondem a uma sequência de decisões binárias aplicadas às variáveis de entrada individual (ver Figura 7.6). Os modelos lineares para a classificação criam superfícies de decisão que são funções lineares do vector de características de entrada e, portanto, são definidas por hiperplanos no espaço de entrada. A superfície de decisão é, portanto, uma simples soma ponderada dos vectores de características de entrada, com pesos maiores sendo atribuídos a características que são mais importantes para o processo de decisão. Em *data mining* e *machine learning*, a árvore da decisão é um modelo preditivo. Nomes mais descritivos para tais modelos da árvore são árvore de classificação (resultado discreto) ou árvore de regressão (resultado contínuo).

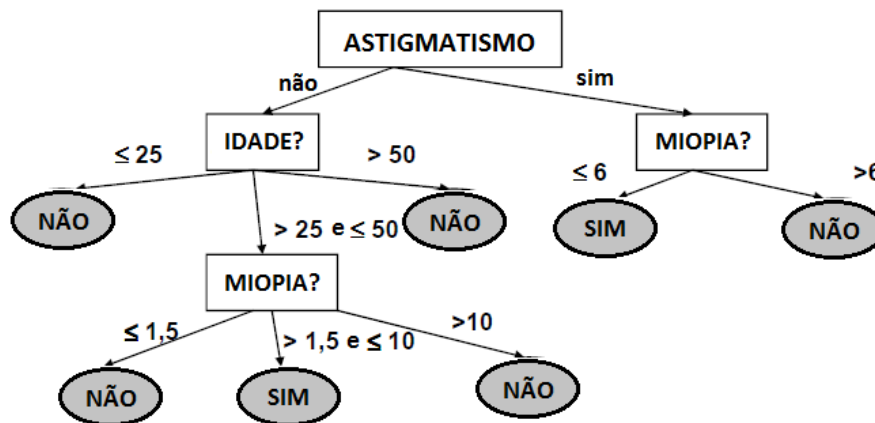


Figura 7.6: Exemplo da estrutura básica de uma árvore de decisão.

Os nós quadrados representam decisões e os nós redondos, nós de incerteza, representam eventos aleatórios. Nos ramos de uma árvore de decisão devem ser anotados:

- as probabilidades após os nós de incerteza;
- os valores de investimentos nos nós de decisão;
- os retornos no final dos ramos;

Para a implementação deste método de classificação foram utilizadas duas plataformas, o WEKA e o MatLab. O WEKA (*Waikato Environment for Knowledge Analysis*¹) é uma ferramenta de código-aberto bastante utilizada. Esta ferramenta, desenvolvida na Universidade de Waikato de Hamilton, Nova Zelândia, implementa mais de vinte algoritmos diferentes de aprendizagem de

¹<http://www.cs.waikato.ac.nz/ml/weka/>

dados convencionais. No MatLab, para a construção de uma árvore de classificação foi utilizada a função já disponibilizada, *classregtree()*. A função referida cria uma árvore binária de decisão para predição da resposta Y em função de dados de treino X . O critério de divisão dos nós aqui utilizado foi o de *Gini*.

Utilizando então as duas plataformas acima referidas procedeu-se à implementação e teste das respectivas árvores de decisão. Os respectivos resultados podem ser observados na Tabela 7.6.

Tabela 7.6: Resultados de classificação utilizando árvores de decisão.

<i>Conjunto Utilizado</i>	<i>WEKA (%)</i>	<i>MatLab (%)</i>
Conjunto fixo	47.50	46.30

Facilmente se conclui que os resultados obtidos a partir deste método não são os melhores. Ainda assim tentou implementar-se um método de rejeição, tal como nos métodos anteriormente descritos, mas sem sucesso. Tal como acontece no SVM, as árvores de decisão não permitem uma saída na qual se possa aplicar um *threshold* de aceitação e dessa forma determinar as taxas FAR e FRR deste método de classificação.

7.3 Conclusão

Chegado à última fase deste sistema, procedeu-se então à implementação e teste de diversos meios para a classificação. Foram implementados e testados três diferentes metodologias possíveis para a fase da identificação. O primeiro, *feature matching* apresenta bons resultados, na ordem dos 73% de GAR e com apenas 8% de FAR. Um outro método utilizado, desta vez de aprendizagem automática, foram os SVMs multi-classe. Este método mostrou-se também bastante eficaz apresentando uma GAR média de 85%. Contudo, devido às suas características, não foram possíveis testes para determinar a FAR que este método apresenta. O último método implementado foram as árvores de decisão. Para tal, recorreu-se a duas plataformas que possibilitam a construção e teste deste tipo de métodos, o WEKA e o MatLab. No entanto, em ambas as plataformas os resultados não foram tão positivos quanto os dos restantes métodos. Aqui obteve-se apenas uma GAR de 48%, não sendo também possível testar a taxa de intrusão, FAR.

Embora não tenha sido mencionado antes, foi também aqui testado nos três métodos o conjunto de dados normalizado e a ordem de importância obtida a partir da função *varimportance()*. Relativamente ao conjunto de dados normalizado, verificou-se que os resultados produzidos por este foram sempre bastante similares aos obtidos com o conjunto de dados com os valores reais e decidiu-se, por isso, a utilização única do segundo. Relativamente à ordem de importância dos atributos obtida, recorrendo à função referida, esta ordem apresentou resultados bastante inferiores

às restantes ordens dos atributos, em todos os testes. Acredita-se que tais resultados menos positivos estejam relacionados com os também menos bons resultados obtidos pelas árvores de decisão, já que a função utilizada recorre à construção de uma árvore binária para cálculo da importância de cada atributo.

Depois de implementados e testados estes três métodos, facilmente se chega à conclusão de que os dois primeiros são aqueles que apresentam melhores resultados, o *feature matching* e os SVMs. No entanto, uma comparação entre estes dois métodos não seria muito realista, dado não se ter conhecimento da FAR obtida no SVM. Ainda assim crê-se que estes dois métodos apresentem resultados bastante semelhantes, mostrando-se assim dois fortes candidatos à aplicação neste tipo de sistemas.

Capítulo 8

Conclusão Geral

Os objectivos iniciais deste trabalho baseavam-se no estudo, implementação e melhoramento das técnicas existentes aplicadas em sistemas de reconhecimento biométrico, mais propriamente, no reconhecimento baseado na geometria da mão. Com o avanço do estudo destes sistemas e revisão da literatura, foi-se percebendo as principais limitações destes sistemas e os objectivos foram, de certa forma, adaptando-se a tais necessidades. Contudo, o objectivo principal focou-se sempre na implementação de um sistema que utilizasse as características geométricas da mão para reconhecimento biométrico, enquanto os objectivos derivados se baseavam em superar as limitações encontradas em sistemas já existentes. Tais limitações residiam, principalmente, nas restrições impostas na fase de aquisição (colocação da mão para aquisição das imagens), desempenho do sistema e selecção/precisão das características extraídas.

Como conclusão geral deste trabalho pode dizer-se que um sistema de reconhecimento biométrico foi aqui desenvolvido desde a fase de aquisição até à fase de identificação. Foi também criada uma BD com imagens da mão sem qualquer tipo de restrições durante a aquisição, contendo dez imagens da mão de 46 indivíduos diferentes.

Relativamente às limitações apresentadas no início deste trabalho na fase de aquisição, nem todas foram superadas. Para a maior parte das imagens o algoritmo de binarização apresenta bons resultados. No entanto, para outras imagens, nomeadamente aquelas em que a mão se encontra com os dedos juntos, o resultado não corresponde ao esperado, limitando assim a BD. Contudo, pode afirmar-se que o sistema desenvolvido apresenta um número bastante menor de restrições na fase de aquisição, o que torna o sistema muito mais confortável e com maior facilidade de interacção.

Na detecção dos pontos característicos vários métodos foram testados apresentando também óptimos resultados, tendo em consideração a redução da BD face a alguns dos resultados menos desejáveis obtidos na etapa de binarização.

Quanto à fase do reconhecimento, foi possível uma boa comparação entre os diferentes métodos de identificação. Também os resultados obtidos para a GAR (72.50%) e a FAR (8.70%) no

método de *feature matching* não se encontram muito distantes do estado da arte, tendo em conta todas as restrições eliminadas na aquisição.

Todos os métodos e algoritmos testados neste trabalho servirão como uma grande base para futuros investigadores nesta área. Mesmo aqueles que apresentaram resultados negativos servirão como base de experiência e teste para futuros desenvolvimentos. Espera-se, também, que a BD criada fomente a investigação nesta área.

Capítulo 9

Trabalho Futuro

Considerando os objectivos inicialmente propostos e os resultados alcançados, surgem naturalmente alguns tópicos para trabalho futuro.

Devem ser investidos esforços na etapa da binarização. Uma possível solução será o desenvolvimento de um algoritmo de segmentação binária adaptado unicamente à segmentação de imagens da mão. Isto é, implementar um método com conhecimento à *priori* do tipo de imagem a tratar de forma a que este apenas termine quando uma correcta e clara segmentação da mão é conseguida.

Apesar dos bons resultados obtidos, há ainda lugar para melhoria dos algoritmos de detecção de pontos. Embora o algoritmo implementado ofereça bons resultados, acredita-se que o método de transformação para coordenadas polares e posterior análise da curvatura seja mais consistente. O grande entrave encontrado neste método foi ao nível da transformação de coordenadas, em que a imagem final não resultava binária e o contorno em coordenadas polares não se encontrava nas condições necessárias para a aplicação do algoritmo de seguimento de contornos. Um melhoramento deste passo levaria a uma simples aplicação da análise da curvatura, que se mostrou bastante eficaz quando o seguimento do contorno é feito com sucesso, tornando a detecção de pontos mais consistente e precisa. Note-se que uma identificação mais precisa dos pontos característicos leva a uma extracção de características também mais realista, melhorando assim todos os resultados finais do sistema.

Algum desenvolvimento poderá também ser considerado na fase de identificação. No método de *feature matching* um algoritmo de atribuição de pesos aos diferentes atributos pode ser realizado de forma a dar mais peso aos atributos com mais importância para a classificação e, dessa forma, influenciar uma melhoria dos resultados. Podem também ser investigadas formas de aplicar um método de rejeição nos classificadores SVM e na árvore de decisão para se conseguirem resultados mais completos e informativos do funcionamento do sistema.

Finalmente, pode ser feito um alargamento da BD criada. Esta BD conta, até ao momento, com imagens de 46 indivíduos, dez por cada um. Pensa-se que o número de imagens por pessoa é suficiente e, embora o número total de pessoas representadas na BD seja suficiente para futuros

estudos e investigações, um aumento da mesma poderá também ser considerado. Esta ampliação deve seguir as condições estabelecidas no início da criação desta BD, de forma a que a BD final siga os mesmos princípios adoptados aquando da sua criação.

Anexo A

Hand Images Data Base

This hand image database contains 460 images captured from 46 individuals, 10 images from each individual.

These images correspond to the right hand of individuals of both sexes, with ages between 10 and 75 years and of various nationalities (Portuguese, Chinese, Indian, Cape Verdean, Brazilian, Polish, Nigerian, Iranian, etc.).

Usually, there are many restrictions on the acquisition method, that in some way affect the feature extraction and the acquisition convenience. In many hand recognition systems it is also common the use of peg's upon image acquisition. This method causes deformation on hand and afterwards the features extracted are not genuine. In the meantime, recent publications use a free peg's acquisition system, but still these systems exhibit some constraints related with hand positioning and orientation as forcing the user to keep their hands opened and fingers spread.

This database was created to develop a revolutionary biometric recognition system, therefore, it has no constraints in hand positioning, orientation or finger arrangement.

All the images were acquired with a normal flatbed scanner, the EPSON PERFECTION V300 PHOTO.

Image Characteristics:

- Resolution: 200 ppp
- Format: .tif
- Color: 24 bits

During the acquisition there are no peg's to restrict hand position and, the user is free to put his/her hand in any possible way, as shown in the next image.



Figura A.1: EPSON PERFECTION V300 PHOTO.

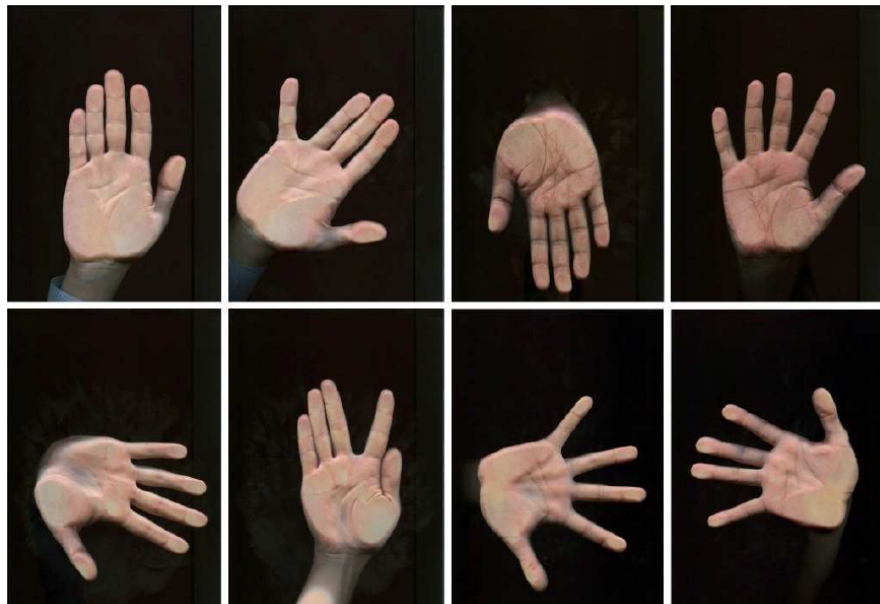


Figura A.2: 8 examples of images in the database. Please notice the freedom given to the individual on the hand positioning and finger arrangement.

Images are stored in different folders, each folder corresponding to a different individual and inside each folder images are numbered from 1-10.

Anexo B

Hand Characteristic Points Detection Competition

Biometrics is one of the most studied subject for purposes of security and identification. Among the various biometrics modalities, such as fingerprint, iris or face recognition, we highlight hand-geometry because of its simplicity in acquisition, for being a non intrusive method and for its wide range of features that can be extracted, from an anatomical point of view. However, there are still many restrictions on the acquisition method, that in some way affect the feature extraction and the commodity on acquisition. In many hand recognition systems is common the use of peg's upon image acquisition (figure 1). This method causes deformation on hand and afterwards the extracted features are not genuine. In the meantime, recent publications use a free peg's acquisition system, but still these systems have some constraints in hand positioning and orientation as forcing the user to keep their hands opened (Figure B.1).



Figura B.1: Acquisition with peg's.



Figura B.2: Examples of restrictions avoided in our Database. a) big part of the arm in image; b) orientation free; c) hand closed.

The most recent papers expose some techniques to extract characteristic points without the using of peg's and with fewer restrictions on positioning, however, as can be seen in Jugurta M. et. al work it's still needed a manual pre-processing of the image to extract these points. In situations where hand images are not ideal the algorithms almost always fail, either because there's a big part of the arm in image or because of the hand being closed, etc.. [1][2][3] The challenge presented on this competition is the development of an algorithm that should be optimized to maximize accurate detection of the 9 characteristic points in hand: 5 finger tips and 4 valleys (Figure B.3). Subsequently, all algorithms will be tested on our own database (Figure 3.2).

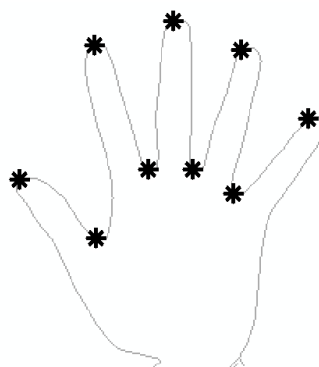


Figura B.3: Characteristic Points to be detected.

The goal of the competition is to compare different methodologies for software-based hand characteristic points detection within a data set composed by real hand images. The ambition of the competition is to become the reference event for academic and industrial researches. It is expected a strong contribution to the state of the art of this particular subject, as well as an improvement in the reliability of this crucial step in hand-geometry based recognition systems.

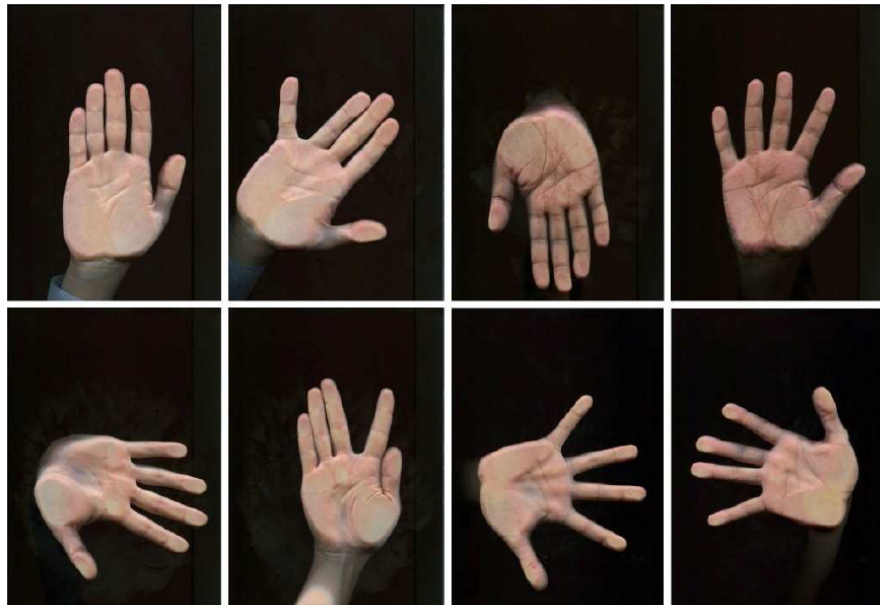


Figura B.4: Some samples of images in our Database.

[1] Nicolas Tsapatsoulis and Constantinos Pattichis, "Palm geometry biometrics: A score-based fusion Approach"

[2] Arun Ross, "A Prototype Hand Geometry-based Verification System"

[3] Jugurta Montalvão, Lucas Molina, Jânio Canuto, "Robust hand image processing for biometric application"

Anexo C

IJUP - Abstract

Hand Geometry Based Recognition System - A New Method with no Constraints on Image Acquisition

Hélder Matos, Filipe Magalhães, Hélder P. Oliveira, Jaime S. Cardoso

*INESC Porto and Faculdade de Engenharia da Universidade do Porto, Rua Dr. Roberto Frias,
s/n 4200 465, Porto, Portugal*

Hand recognition systems are among the oldest biometric tools used for automatic person authentication. Access control devices have been manufactured and commercialized since the late 1970s. Several patents have already been issued for hand recognition devices and live applications have been launched and used in nuclear plants, airports and hotels over the last 30 years. Hand-based recognition is reliable, low-cost, intuitive and non-invasive. It is a viable solution for a range of access control applications.

However, recognition systems based on hand geometry are still far from the favouritism that other types of systems already possess. This unpopularity is related with the restrictions imposed on acquisition of images to use on the Recognition or Identification phase. Despite of the performances stated by the state-of-the-art, there are still many limitations on the hand positioning upon the acquisition. In some systems pegs are used to force the user to place his hand in a certain position [1]. However, in addition to being uncomfortable, these pegs cause deformations on hand, which subsequently will influence the feature extraction. Other systems use a peg-free methodology [2], which allows a little more of freedom to the user, but at the same time it is imposed to the user a predetermined hand position.

In this work we propose to design a system capable of showing an acceptable performance for a bigger universe of users and also eliminate all or, at least, most of the restrictions imposed on the acquisition. The exclusion of restrictions on the acquisition will enable a significant advance evolution of these systems, mainly because they will become simpler to use, with a considerable improvement in the interaction with user.

Thus, we developed a new algorithm that uses the skeleton of the hand binary image to detect the 5 fingertip's points. For detection of the 4 valleys between fingers, this algorithm uses an edge following algorithm. This method has reached a rate of 97% accuracy on the detection of the desired points, and was applied to images without restrictions on the acquisition. After point detection, 21 features are extracted and then a recognition algorithm, based on feature-matching, is applied. This system presents a Genuine Acceptance Rate of 73% with a False Acceptance Rate of only 8%.

Within this work, a new image database was also created with images acquired without restrictions to the hand positioning. Besides being public and without antecedents, this database also allows a greater motivation on the study of such systems and their constraints.

References:

- [1] Marcos Faundez-Zanuy and Guillermo Mar Navarro Mérida (2005), Biometric Identification by Means of Hand Geometry, Computational Intelligence and Bioinspired Systems and a Neural Net Classifier, *Personal Psychology*, 22 (2), 391-413.
- [2] Jugurta Montalvão, Lucas Molina and Jânio Canuto (2010), Robust hand image processing for biometric application, *Pattern Analysis & Applications*, Volume 13, Number 4, 397-407.

Referências

- Miguel Adan, Antonio Adan, Andres S. Vazquez e Roberto Torres. Biometric verification/identification based on hands natural layout. *Image and Vision Computing*, 26(4):451 – 465, 2008. ISSN 0262-8856.
- Theodosiou Z. Antonakoudi A., Kounoudes A. Polybio multibiometrics database: Contents, description and interfacing. *Proceedings of the Workshop on Artificial Intelligence Approaches for Biometric Template Creation and Multibiometrics Fusion (ArtIBio)*, páginas 1168 – 1171, 2009.
- Christopher M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006. ISBN 0387310738.
- Chih-Chung Chang e Chih-Jen Lin. *LIBSVM: a library for support vector machines*, 2001. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- Marcos Faundez-Zanuy e Guillermo Mar Navarro Merida. Biometric identification by means of hand geometry and a neural net classifier. In Joan Cabestany, Alberto Prieto e Francisco Sandoval, editors, *Computational Intelligence and Bioinspired Systems*, volume 3512 of *Lecture Notes in Computer Science*, páginas 1172–1179. Springer Berlin / Heidelberg, 2005.
- Marcos Faundez-Zanuy, David A. Elizondo, Miguel Ángel Ferrer-Ballester e Carlos M. Travieso-González. Authentication of individuals using hand geometry biometrics: A neural network approach. *Neural Process Lett*, páginas 201–216, 2007.
- R. E. Woods Gonzalez, R. C. e S. L. Eddins. *Digital image processing using matlab*. 2004.
- A.K. Jain e A. Ross. Multibiometric systems. *Communications of the ACM*, páginas 34 – 40, 2004.
- A.K. Jain, L. Hong e S. Pankanti. Biometric identification. *Communications of the ACM*, páginas 91 – 98, 2000.
- Anastasis Kounoudes, Nicolas Tsapatsoulis, Zenonas Theodosiou e Marios Milis. Polybio: Multimodal biometric data acquisition platform and security system. In Ben Schouten, Niels Juul, Andrzej Drygajlo e Massimo Tistarelli, editors, *Biometrics and Identity Management*, volume 5372 of *Lecture Notes in Computer Science*, páginas 216–227. Springer Berlin / Heidelberg, 2008.
- P. D. Kovesi. *MATLAB and Octave functions for computer vision and image processing*. Available from: <<http://www.csse.uwa.edu.au/~pk/research/matlabfns/>>.

- Ajay Kumar, David Wong, Helen Shen e Anil Jain. Personal verification using palmprint and hand geometry biometric. In Josef Kittler e Mark Nixon, editors, *Audio- and Video-Based Biometric Person Authentication*, volume 2688, páginas 1060–1060. Springer Berlin / Heidelberg, 2003.
- Dongju Liu e Jian Yu. Otsu method and k-means. In *Hybrid Intelligent Systems, 2009. HIS '09. Ninth International Conference on*, volume 1, páginas 344–349, 2009.
- Filipe Magalhaes, Helder Oliveira e Aurelio Campilho. A new method for the detection of singular points in fingerprint images. *Proceedings of the IEEE Workshop on Applications of Computer Vision*, páginas 157–162, 2009.
- Jugurta Montalvao, Lucas Molina e Janio Canuto. Robust hand image processing for biometric application. *Pattern Analysis and Applications*, 13:397–407, 2010. ISSN 1433-7541.
- S. Prabhakar, S. Pankanti e A.K. Jain. Biometric recognition: security and privacy concerns. *Security Privacy, IEEE*, páginas 33 – 42, 2003.
- Ratika Pradhan, Shikhar Kumar, Ruchika Agarwal, Mohan P. Pradhan e M. K. Ghose. Contour line tracing algorithm for digital topographic maps. *International Journal of Image Processing (IJIP)*, 2010.
- A. Ross, S. Pankanti e A.K. Jain. A prototype hand geometry-based verification system. *Proceedings of 2nd Conference on Audio and Video based Biometric Person Authentication (AVBPA)*, páginas 166 – 171, 1999.
- Vitomir Struc e Nikola Pavesic. Hand-geometry device. *Encyclopedia of Biometrics*, páginas 693–698, 2009.
- Tsapatsoulis e C.S. Pattichis. Palm geometry biometrics: A score-based fusion approach. *Proceedings of AIAI*, 425:158–167, 2009.
- Ra L. N. Wong e Pengcheng Shi. Peg-free hand geometry recognition using hierarchical geometry and shape matching. In *IAPR Workshop on Machine Vision Applications*, páginas 281–284, 2002.