

Faculdade de Ciências e Faculdade de Engenharia da Universidade do Porto

# AUTENTICAÇÃO DE FACES A PARTIR DA AQUISIÇÃO DE SEQUÊNCIAS DE IMAGENS

Ana Raquel Ferreira de Almeida Sebastião

Licenciada em Matemática – ramo Matemática Aplicada, pela Faculdade de  
Ciências da Universidade do Porto

Dissertação provisória submetida para efeitos de atribuição do grau de Mestre em  
Métodos Computacionais em Ciências e Engenharia

Sob orientação científica do Doutor Armando Jorge Monteiro Neves Padilha,  
Professor Associado da Faculdade de Engenharia da Universidade do Porto  
e do Doutor Jorge Alves da Silva,  
Professor Auxiliar da Faculdade de Engenharia da Universidade do Porto

Porto, Janeiro de 2006



# Resumo

Inserida no domínio da visão por computador, esta dissertação teve por objectivo a construção de um sistema automático de visão, permitindo a aquisição, o processamento e o reconhecimento de informação visual relativa à face humana, para efeitos de autenticação de “clientes”. A principal finalidade deste processo de autenticação é controlar de forma automática o acesso a áreas restritas, evitando a entrada de “impostores”.

O trabalho de investigação efectuado teve em vista analisar a possibilidade de mudança de paradigma relativamente aos sistemas de identificação e autenticação de faces humanas, com a finalidade de proporcionar maior conforto aos “clientes” do sistema, quer na fase de inserção na base de dados, quer na fase de teste (acesso às instalações). Assim, em vez de se investir na criação de uma base de dados constituída por várias (às vezes, numerosas) vistas da face de cada indivíduo, correspondentes a diversas poses, diferentes estados emocionais e uso de distintos adereços, com vista a potenciar o grau de reconhecimento do indivíduo a partir da captura de uma única imagem de teste, o trabalho realizado “inverteu” o paradigma, criando uma base de dados com apenas uma vista frontal de cada indivíduo e usando múltiplas imagens de teste extraídas de uma sequência de vídeo captada durante a aproximação do indivíduo à câmara.

Apresentam-se alguns dos métodos passíveis de utilização no reconhecimento de pessoas por recurso à análise de imagens, sendo adoptada a técnica de Análise de Componentes Independentes (ICA, acrónimo derivado da terminologia anglo-saxónica).

As imagens utilizadas para testar a autenticação foram retiradas de sequências de vídeo, praticamente não condicionadas e de qualidade relativamente fraca. Por as pessoas não se apresentarem sempre em pose frontal perante a câmara, foi incorporada no trabalho a síntese de vistas frontais, recorrendo-se a uma técnica de *View-Morphing*.

Com o objectivo de melhorar a qualidade e a taxa de reconhecimento foi efectuado um pré-processamento das imagens, utilizando uma transformação fotométrica e uma normalização geométrica que permitiram compensar pequenas distorções geométricas e diminuir o efeito das variações da luminosidade.

Os resultados obtidos comprovam que a utilização de pré-processamento nas imagens conduz a melhores resultados, que se reflectem num aumento da taxa de acerto. A síntese de vistas frontais, através da técnica de *View-Morphing*, revela resultados de identificação melhores do que os obtidos com as imagens extraídas directamente da sequência vídeo. Para testar a qualidade e a eficácia da autenticação foram efectuados testes do tipo “*Leave-One-Out*”, que apresentam resultados animadores, mas que indiciam que o condicionamento do ambiente para a aquisição das imagens da base de dados, bem como as da sequência vídeo, deve ser mais detalhadamente analisado.

# Abstract

As a contribution to the computer vision research area, this dissertation intends to construct an automatic vision system, allowing the acquisition, the processing and the recognition of human face visual information, in order to authenticate "clients". The main purpose of this authentication process is to control in an automatic way the access to restricted areas, avoiding the entrance of "impostors".

The development done allows us to analyze the possibility of changing the paradigm that concerns the identification and authentication systems of human faces, providing a comfortable situation to "clients" in both in both the database insertion and the test phases (facilities access). The current standard is that of creating a database with several (sometimes, numerous) views of each individual's face, corresponding to distinct poses, different emotional states and the use of some accessories such as glasses, earrings and necklaces; these multiple views are intended to increase the individual's recognition rate for the capture of a single test image. This project "reversed" the paradigm, by creating a database with a single front view of each individual and by using multiple test images extracted from a video sequence captured during the individual's approach path to the camera.

Some of the methods susceptible to be used in face recognition based on image analysis are described in this document. We have chosen to use the Independent Components Analysis (ICA) technique.

The images used for authentication were selected from video sequences, practically free of restrictions and with relatively poor quality. Because subjects didn't always present themselves in a front pose before the camera, the synthesis of front views was incorporated, by using a View-Morphing method.

Another goal of this work was to improve the quality and the recognition rate through image pre-processing, using a photometric transformation as well as a geometric normalization that allowed to compensate for small geometric distortions and to reduce the effect of brightness variations.

The obtained results prove that pre-processing the images leads to better results, increasing the success rate. The synthesis of front views, through the View-morphing technique, reveals better identification results than the ones obtained with the images directly selected from the video sequence. To test the quality and the effectiveness of the authentication, "*Leave-One-Out*" tests were done, providing promising results, but nevertheless showing that the environment conditioning for the acquisition of the database images, as well as for the video sequence, should be improved and carefully analyzed.

# Résumé

Insérée dans le domaine de la vision par ordinateur, cette dissertation a eu pour objectif la construction d'un système automatique de vision, permettant l'acquisition, le traitement et la reconnaissance d'information visuelle relative à la face humaine, en ce qui concerne l'authentification "de clients". La principale finalité de ce processus d'authentification est de contrôler de façon automatique l'accès à des aires restreintes, évitant l'entrée de "trompeurs".

Le développement effectué a eu pour objectif analyser la possibilité de changement de paradigme par rapport aux systèmes d'identification et authentification de faces humaines, ayant la finalité de proportionner plus de confort aux "clients" du système, soit dans la phase d'insertion dans la base de données, soit dans la phase de test (accès aux installations). Ainsi, au lieu de s'investir dans la création d'une base de données constituée par plusieurs (quelquefois, nombreuses) vues de chaque individu, correspondantes à diverses positions, différents états émotionnels et l'usage de différentes parures, ayant pour objectif augmenter le degré d'identification de l'individu à partir de la capture d'une seule image d'essai, le travail réalisé "a inversé" le paradigme, créant une base de données ayant seulement une pose de face de chaque individu et en utilisant multiples images de test extraites d'une séquence de vidéo captée pendant l'approximation de l'individu à la caméra.

Après la présentation de quelques méthodes passibles d'être utilisées pour la reconnaissance de personne dans le recours à l'analyse d'images, suit la technique de l'Analyse en Composantes Indépendantes (du Anglais ICA – *Independent Component Analysis*).

Les images utilisées pour tester l'authentification ont été retirées de séquences vidéo, non conditionnées et de qualité relativement faible. Dans la mesure où elles ne présentent pas toujours une pose de face par rapport au centre optique de la caméra, il a été nécessaire de synthétiser des vues de face en ayant recours à la technique *View-Morphing*.

Dans l'objectif d'améliorer la qualité et les taux de reconnaissance, un pré-traitement des images a été effectué, en utilisant une transformation photométrique et une normalisation géométrique qui ont permis de corriger des petites distorsions géométriques ainsi que de diminuer l'effet des variations de luminosité.

Les résultats obtenus prouvent que l'utilisation du pré-traitement des images conduit à de meilleurs résultats, qui se reflètent dans l'augmentation du taux de réussite. La synthèse de vues de face, à travers la technique *View-Morphing*, révèle des résultats d'authentification meilleurs par rapport aux résultats obtenus avec les images retirées directement de la séquence vidéo. Pour tester la qualité et l'efficacité de l'authentification des tests "*Leave-One-Out*" ont été conduits, qui présentent des résultats encourageants, mais qui accusent que les circonstances de l'environnement pour l'acquisition des images de la base de données, aussi bien que celles de la séquence vidéo, doit être analysé plus en détail.

# Agradecimentos

Começo por agradecer ao Professor Armando Jorge Padilha não só pela sua orientação científica como também pela arte em transmitir conhecimentos, pelo encorajamento nos momentos de desânimo e pela ponderação e disponibilidade manifestadas no decurso da minha dissertação.

Ao Professor Jorge Alves da Silva agradeço pelo apoio e supervisão, pelo trabalho de revisão da presente dissertação e ainda pelos conhecimentos transmitidos.

Aos colegas de curso e de trabalho, obrigada pelo vosso apoio moral, pelo sentido crítico e pelos incentivos.

Aos amigos de sempre, que me acompanharam no decurso deste trabalho, obrigada por demonstrarem a vossa amizade nas horas difíceis e pela força que me deram.

À minha família, em especial aos meus pais e à minha irmã, que há tanto anseiam por este momento, agradeço o amor com que me acompanharam neste percurso, a compreensão, o apoio e a paciência desmedida.

Ao Bruno, meu melhor amigo e companheiro, quero agradecer pelo amor e carinho, pelo apoio inestimável por acreditar em mim e por me mostrar a luz nos momentos em que me faltavam as forças e ânimo. Obrigada por estares sempre comigo!

Por fim, agradeço a todos aqueles que disponibilizaram a sua imagem para a construção da base de dados efectuada neste trabalho.

# Índice

|       |                                                        |    |
|-------|--------------------------------------------------------|----|
| 1     | Introdução .....                                       | 1  |
| 1.1   | Motivação e Objectivos .....                           | 2  |
| 1.2   | Organização da Dissertação .....                       | 3  |
| 2     | Análise de Imagens para Reconhecimento de Pessoas..... | 5  |
| 2.1   | Revisão dos Métodos Utilizados .....                   | 5  |
| 2.2   | Representação de Imagens .....                         | 8  |
| 2.3   | Análise de Factores .....                              | 10 |
| 2.3.1 | Análise de Componentes Principais .....                | 11 |
| 2.3.2 | Análise de Factores Comuns .....                       | 13 |
| 2.4   | Análise de Componentes Independentes .....             | 14 |
| 2.4.1 | Métodos de Estimação .....                             | 18 |
| 2.4.2 | Algoritmos de Optimização .....                        | 23 |
| 2.4.3 | Representação Matricial dos Dados.....                 | 24 |
| 3     | Síntese de Vistas Frontais.....                        | 28 |
| 3.1   | Revisão do Estado da Arte .....                        | 28 |
| 3.2   | View-Morphing .....                                    | 31 |
| 3.2.1 | Vistas Paralelas .....                                 | 33 |
| 3.2.2 | Vistas Não Paralelas .....                             | 36 |
| 3.3   | Implementação da técnica <i>View-Morphing</i> .....    | 38 |
| 3.3.1 | <i>Prewarp</i> .....                                   | 38 |
| 3.3.2 | <i>Morph</i> .....                                     | 42 |
| 3.3.3 | <i>Postwarp</i> .....                                  | 43 |
| 4     | Autenticação através de ICA .....                      | 45 |

|       |                                                                                                                   |    |
|-------|-------------------------------------------------------------------------------------------------------------------|----|
| 4.1   | Construção da Base de Dados .....                                                                                 | 45 |
| 4.2   | Síntese de Vistas Frontais das Imagens seleccionadas a partir da<br>Sequência Vídeo .....                         | 47 |
| 4.3   | Pré-Processamento das Imagens .....                                                                               | 51 |
| 4.3.1 | Transformação Geométrica .....                                                                                    | 52 |
| 4.3.2 | Transformação Fotométrica.....                                                                                    | 53 |
| 4.4   | Processamento dos Dados .....                                                                                     | 56 |
| 4.4.1 | Pré-Processamento.....                                                                                            | 56 |
| 4.4.2 | Processamento - Selecção do Método de Estimação, Algoritmo e<br>Função Objectivo Utilizados na Autenticação ..... | 57 |
| 4.4.3 | Pós-Processamento: Escolha da medida de distância entre imagens.....                                              | 59 |
| 5     | Resultados Obtidos .....                                                                                          | 61 |
| 5.1   | Comparação entre os Resultados de Identificação das Imagens da<br>Sequência Vídeo e das Imagens Sintetizadas..... | 65 |
| 5.2   | Comparação entre os Resultados Sem e Com Correção Geométrica<br>das Imagens .....                                 | 69 |
| 5.2.1 | Redução da Dimensionalidade dos Dados.....                                                                        | 73 |
| 5.3   | Autenticação Utilizando as Vistas Frontais Sintetizadas Com Pré-<br>Processamento .....                           | 75 |
| 5.4   | Testes de Rejeição de Impostores .....                                                                            | 77 |
| 6     | Conclusões e Perspectivas Futuras .....                                                                           | 81 |
| 6.1   | Conclusões.....                                                                                                   | 81 |
| 6.2   | Perspectivas Futuras .....                                                                                        | 83 |
|       | Referências .....                                                                                                 | 85 |
|       | Apêndice A - Notação.....                                                                                         | 89 |

# Lista de Figuras

|                                                                                                                                                                                                                                                                                                                                                                                                   |    |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figura 2.1 - Esquema de Reconhecimento de Imagens.....                                                                                                                                                                                                                                                                                                                                            | 7  |
| Figura 2.2 - Modelo de Estimação das Componentes Independentes. Adaptado de [Bartlett_98].....                                                                                                                                                                                                                                                                                                    | 15 |
| Figura 2.3 - Modelo de separação das componentes independentes. Adaptado de [Bartlett_98].....                                                                                                                                                                                                                                                                                                    | 16 |
| Figura 2.4 - Representação das faces como combinação linear das componentes independentes. Adaptado de [Bartlett_98].....                                                                                                                                                                                                                                                                         | 16 |
| Figura 2.5 - Separação das fontes da imagem obtendo as componentes independentes nas linhas de $\hat{\mathbf{s}}$ . Adaptado de [Bartlett_98]. ....                                                                                                                                                                                                                                               | 25 |
| Figura 2.6 - Separação das fontes dos <i>pixels</i> obtendo factorial <i>face code</i> nas colunas de $\hat{\mathbf{s}}$ . Adaptado de [Bartlett_98]. ....                                                                                                                                                                                                                                        | 26 |
| Figura 3.1 - Exemplo de Síntese de Vistas Intermédias, adaptado de [Seitz_97]. ....                                                                                                                                                                                                                                                                                                               | 28 |
| Figura 3.2 - Síntese de imagem através da técnica <i>Image Morphing</i> entre uma vista de Mona Lisa e a sua reflexão, provocando contracção horizontal e vertical da face e do busto, [Seitz_97].....                                                                                                                                                                                            | 31 |
| Figura 3.3 - Distorção provocada na síntese de imagens através da técnica <i>Image Morphing</i> . Ilustra-se uma interpolação linear entre duas vistas de um relógio (imagem mais à esquerda e mais à direita) que provoca um efeito de curvatura nas imagens mais intermédias. A linha a tracejado representa o percurso de uma característica em diversas imagens intermédias, [Seitz_97]. .... | 32 |
| Figura 3.4 - Síntese de imagem através da técnica <i>View Morphing</i> , entre uma vista da Mona Lisa e a sua reflexão, produzindo uma vista mais realística, [Seitz_97]. ....                                                                                                                                                                                                                    | 32 |
| Figura 3.5 - Síntese da imagem intermédia $I_{0.5}$ , entre duas vistas distintas do autocarro ( $I_0$ e $I_1$ ), através da técnica <i>View-Morphing</i> , [Seitz_97]. ....                                                                                                                                                                                                                      | 33 |
| Figura 3.6 – <i>View-Morphing</i> de vistas paralelas. A interpolação linear de <i>pixels</i> correspondentes em vistas paralelas com planos de imagem $I_0$ e $I_1$ produz a imagem $I_{0.5}$ que representa outra vista paralela do mesmo objecto ou cena do mundo 3D, [Seitz_97].....                                                                                                          | 35 |
| Figura 3.7 - <i>View-Morphing</i> de vistas não paralelas. As imagens originais $I_0$ e $I_1$ são <i>prewarped</i> de modo a formarem as imagens $\hat{I}_0$ e $\hat{I}_1$ . A imagem $\hat{I}_s$ é obtida por interpolação entre as imagens $\hat{I}_0$ e $\hat{I}_1$ . Para finalizar a imagem $\hat{I}_s$ é <i>postwarped</i> , originando $I_s$ , [Seitz_97]. ....                            | 37 |
| Figura 3.8 – Plano $E$ que intersecta $I_0$ e $I_1$ . [Seitz_97]. ....                                                                                                                                                                                                                                                                                                                            | 40 |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figura 3.9 - Exemplo do procedimento da técnica <i>View-Morphing</i> : selecção de um conjunto de características (linhas amarelas) nas imagens $I_0$ e $I_1$ ; utilização destas características para obter as imagens <i>prewarped</i> $\hat{I}_0$ e $\hat{I}_1$ . A imagem intermédia $\hat{I}_{0.5}$ é obtida por interpolação entre as imagens $\hat{I}_0$ e $\hat{I}_1$ . Para finalizar a imagem $\hat{I}_s$ é <i>postwarped</i> , originando $I_s$ , [Seitz_97]..... | 43 |
| Figura 4.1 - Processo de selecção, extracção da região de interesse e redimensionamento das imagens.....                                                                                                                                                                                                                                                                                                                                                                     | 48 |
| Figura 4.2 - Imagens simétricas. ....                                                                                                                                                                                                                                                                                                                                                                                                                                        | 48 |
| Figura 4.3 - Síntese de imagens intermédias, através da técnica <i>View-Morphing</i> , dependente da escolha do parâmetro $s$ .....                                                                                                                                                                                                                                                                                                                                          | 50 |
| Figura 4.4 - Correspondência dos pontos de controlo entre o referencial original (esquerda) e o referencial fixo (direita), Adaptado de [Silva_04]. ....                                                                                                                                                                                                                                                                                                                     | 53 |
| Figura 4.5- Exemplos da aplicação da equalização do histograma. Esquerda - Imagens originais. Direita - Imagens resultado.....                                                                                                                                                                                                                                                                                                                                               | 55 |
| Figura 4.6 - Exemplos de equalização do histograma. Esquerda - Histograma das imagens originais. Centro - Gráfico das transformações aplicadas. Direita - Histograma das imagens resultado. ....                                                                                                                                                                                                                                                                             | 55 |
| Figura 5.1 – Imagens, em pose frontal, de alguns indivíduos que constituem a base de dados.....                                                                                                                                                                                                                                                                                                                                                                              | 63 |
| Figura 5.2 – Imagens, em pose frontal, de alguns indivíduos que constituem o conjunto de treino $\mathbf{X}_{\text{treino}_1}$ . ....                                                                                                                                                                                                                                                                                                                                        | 66 |
| Figura 5.3 – Algumas imagens extraídas da sequência de vídeo e que constituem o conjunto de teste $\mathbf{X}_{\text{teste}_1}$ . ....                                                                                                                                                                                                                                                                                                                                       | 67 |
| Figura 5.4 – Algumas imagens sintetizadas, a partir das imagens extraídas da sequência de vídeo, e que constituem o conjunto de teste $\mathbf{X}_{\text{teste}_2}$ . ....                                                                                                                                                                                                                                                                                                   | 68 |
| Figura 5.5 – Imagens, em pose frontal, de alguns indivíduos que constituem o conjunto de treino $\mathbf{X}_{\text{treino}_2}$ . ....                                                                                                                                                                                                                                                                                                                                        | 70 |
| Figura 5.6 – Algumas imagens sintetizadas e após correcção geométrica e que constituem o conjunto de teste $\mathbf{X}_{\text{teste}_3}$ . ....                                                                                                                                                                                                                                                                                                                              | 71 |
| Figura 5.7 – Relação entre a taxa de identificação e o número de componentes independentes estimadas, considerando as metodologias com e sem normalização geométrica. ....                                                                                                                                                                                                                                                                                                   | 73 |
| Figura 5.8 – Relação entre a taxa de autenticação e o número de componentes independentes estimadas, considerando as metodologias com e sem normalização geométrica. ....                                                                                                                                                                                                                                                                                                    | 74 |

# Lista de Tabelas

|                                                                                                                                                                                                                             |    |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Tabela 5.1 - Taxas de acerto obtidas para a identificação por aplicação da técnica ICA às imagens retiradas da sequência vídeo e às imagens sintetizadas a partir destas.....                                               | 68 |
| Tabela 5.2 - Taxas de acerto para a identificação de imagens obtidas por aplicação da técnica ICA às imagens sintetizadas sem e com correcção geométrica.....                                                               | 72 |
| Tabela 5.3 - Taxas de acerto para a autenticação de imagens obtidas por aplicação da técnica ICA às imagens sintetizadas sem e com correcção geométrica.....                                                                | 72 |
| Tabela 5.4 - Distâncias das representações ICA das imagens sintetizadas (com equalização do histograma e com correcção geométrica) às representações do conjunto de treino $\mathbf{X}_{\text{treino}_2}$ .....             | 76 |
| Tabela 5.5 - Distâncias das representações ICA das imagens sintetizadas (com pré-processamento) do indivíduo $j$ ( $j=1,2,\dots,22$ ) às representações do conjunto de treino constituído pelos restantes 21 elementos..... | 79 |
| Tabela A.1 - Notação utilizada e respectiva descrição.....                                                                                                                                                                  | 89 |

# 1 INTRODUÇÃO

O reconhecimento de imagens, nomeadamente o reconhecimento de faces humanas por recurso à técnica de Análise de Componentes Independentes (ICA, acrónimo derivado da terminologia anglo-saxónica), tem sido alvo de grande atenção, originando vastos estudos e investigações que evidenciam as suas inúmeras potencialidades e sustentam a sua aplicabilidade nas mais diversas disciplinas científicas, tais como Processamento de Sinal e de Imagem, Problemas de Séries Temporais e Problemas Financeiros, entre outros.

Intimamente ligadas ao reconhecimento de faces surgem a autenticação de faces e a identificação de faces em ambientes de contexto complexo.

A identificação de faces humanas tem por objectivo confrontar uma imagem facial com todas as imagens presentes numa base de dados que é o repositório do universo de reconhecimento. Ao invés, a autenticação visa verificar a correspondência entre uma ou mais vistas da face de um indivíduo, que anuncia a sua identificação, com a representação da pessoa anunciada na base de dados. A principal finalidade do processo de autenticação é controlar de forma automática o acesso a áreas restritas.

Inserido no domínio da visão por computador, neste projecto concebeu-se um sistema automático de visão e autenticação, permitindo a aquisição, processamento e reconhecimento de informação, neste caso, relativa à face humana. A forte expansão sofrida na área de processamento e análise de imagens faciais facilita o acesso a

inúmeros trabalhos e a várias bases de dados testadas e utilizadas para fins de reconhecimento. Estas bases de dados englobam, geralmente, várias imagens faciais do mesmo indivíduo, representando-o em diversas poses, expressões faciais e emocionais, e com a presença de “adereços” como óculos, barba, chapéu, maquilhagem, entre outros. O facto de existirem diferentes representações relativas à mesma face torna o processo de reconhecimento moroso, dificultando assim o controle de acesso a áreas restritas.

Na abordagem deste trabalho visa-se a construção de uma base de dados com apenas uma imagem facial de cada indivíduo, em pose frontal perante a câmara e com uma expressão neutra. A principal vantagem esperada desta construção da base de dados é a diminuição de memória ocupada e do tempo de reconhecimento, permitindo assim proceder à autenticação de clientes em sistemas automáticos de segurança, em tempo real. Não se ignora, contudo, que a redução da base de dados acarretará, por outro lado, uma diminuição da robustez e especificidade das classificações das faces, apesar dos benefícios que poderão ser colhidos pelo facto de, em cada sequência vídeo, estarem disponíveis diversas imagens em posturas diversas.

## **1.1 Motivação e Objectivos**

A motivação para desenvolver este trabalho surgiu, essencialmente, da concretização de um projecto anterior, realizado por Silva, [Silva\_04], na Faculdade de Ciências e Faculdade de Engenharia da Universidade do Porto, com o tema “Reconhecimento de Imagem por meio da Análise de Componentes Independentes”.

Esse trabalho teve como tema principal a identificação de faces humanas, obtidas numa base de dados de domínio público, através da técnica de análise de componentes independentes.

Foi objectivo parcial do presente projecto a construção de uma base de dados contendo apenas uma imagem facial de cada indivíduo, em pose frontal perante a câmara, em tons de cinzento e com fundo o mais uniforme possível.

O procedimento de autenticação foi efectuado através de várias imagens obtidas a partir de uma sequência vídeo. Assim, o foco deste trabalho foi a autenticação de faces e não a identificação das mesmas, visando alcançar os seguintes objectivos:

- Após construção da base de dados, efectuar a normalização geométrica e fotométrica das imagens adquiridas.
- Proceder à síntese de uma vista frontal (para cada imagem seleccionada da sequência vídeo), e submeter estas vistas às transformações referidas anteriormente, para efectuar a autenticação.
- Utilizando a análise de componentes independentes, determinar a distância de cada imagem de teste à imagem correspondente da base de dados, verificando-se autenticação caso a distância mínima não ultrapasse um limiar de significância.

## 1.2 Organização da Dissertação

A presente dissertação encontra-se organizada em seis capítulos, nos quais são abordadas as técnicas utilizadas na realização do projecto de mestrado, análise dos resultados, conclusões e objectivos futuros.

No segundo capítulo, descrevem-se técnicas de representação de imagens, focando em particular a Análise de Componentes Independentes como método implementado para proceder à autenticação.

O capítulo seguinte é dedicado à síntese de novas vistas a partir de duas vistas que representam uma cena ou objecto 3D. Recorreu-se a estes métodos de síntese de vistas intermédias para, a partir das vistas laterais adquiridas por decimação de uma sequência vídeo, criar vistas frontais do indivíduo passíveis de utilizar nos processos de identificação e de autenticação.

A construção da base de dados é abordada no quarto capítulo, referindo as condições criteriosas de captura de imagens representativas da face humana e as transformações fotométricas e geométricas, a utilizar com o objectivo de diminuir as discrepâncias entre as imagens e aumentar a qualidade destas. São também abordados

alguns tipos de pré-processamento de dados, nomeadamente a centralização dos dados e o branqueamento dos mesmos. De seguida, apresentam-se as escolhas do método de estimação, do algoritmo e da função objectivo utilizados na Identificação e na Autenticação, bem como as respectivas justificações. Este capítulo termina com uma definição que permite avaliar a diferença entre representações ICA das imagens da base de dados e das obtidas da sequência vídeo.

No quinto capítulo, apresentam-se e analisam-se os resultados obtidos. É efectuada uma sumarização das questões que o trabalho desenvolvido permitiu responder e abordam-se novas questões que surgiram no decorrer do mesmo.

Por fim, apresentam-se as conclusões a que o trabalho desempenhado conduziu e delineiam-se perspectivas futuras de desenvolvimento, demarcando caminhos a seguir, estando estes sujeitos a possíveis alterações.

## **2 ANÁLISE DE IMAGENS PARA RECONHECIMENTO DE PESSOAS**

Neste capítulo apresentam-se alguns dos métodos passíveis de utilizar no reconhecimento de pessoas por recurso à análise de imagens. Este capítulo inicia-se por uma breve introdução a outras técnicas de reconhecimento que não utilizam a análise de imagens.

### **2.1 Revisão dos Métodos Utilizados**

Nos dias de hoje, os métodos convencionais (como cartões magnéticos, o *login* do conjunto *utilizador/password* e *smart cards*) não garantem a segurança no acesso a

áreas restritas, já que os mesmos podem ser utilizados indevidamente por outros ou esquecidos, perdidos, falsificados e ainda apresentar defeitos.

A Biometria, que é a ciência da identificação baseada na medição precisa de traços biológicos, surge como resposta às limitações apresentadas. De facto, com um *login* biométrico pretende-se conseguir que só indivíduos autorizados o utilizem, evitando assim a intrusão e a falsificação.

No que concerne à segurança, a biometria permite a verificação da identidade de um indivíduo através da mensuração de características fisiológicas únicas inerentes ao mesmo, tais como impressão digital, geometria da palma da mão e seus sulcos, padrão da íris ou da retina do olho humano, reconhecimento facial ou de voz, entre outros.

Os métodos baseados no reconhecimento da íris e da retina são os que, presentemente, conferem maior fiabilidade; contudo, o seu carácter invasivo tem impulsionado estudos e investigações científicas mais profundas em reconhecimento de faces humanas, permitindo a identificação sem incómodo para o indivíduo.

Utilizando técnicas muito variadas, os métodos de reconhecimento de faces, nos últimos tempos, têm captado a atenção e têm sido alvo de diversas pesquisas, sofrendo uma forte expansão intimamente relacionada com as inúmeras aplicações que permitem, das quais se distinguem:

- Seguimento de indivíduos em estádios, estações de metro, etc.;
- Sistemas de segurança e de controlo de acesso a áreas restritas;
- Interfaces perceptuais homem-máquina com reconhecimento de expressões faciais;
- Procuras em ficheiros criminais;
- Envelhecimento computadorizado como meio auxiliar na busca de desaparecidos.

Relacionada com a problemática do reconhecimento de imagens surge a identificação de faces em ambientes complexos, pelo que existem diversas bases de dados resultantes de trabalhos nesta área. No âmbito da identificação de faces, as bases de dados constroem-se com várias imagens representativas, com diferentes poses, diferentes expressões faciais e com a presença de “adereços” como óculos, barba, chapéu, entre outros. Para efectuar identificação, confronta-se apenas uma imagem de um indivíduo (captada no momento de acesso) com as várias imagens da base de dados.

Neste trabalho construiu-se uma base de dados com características diferentes das usualmente utilizadas para o efeito. A base de dados construída contém apenas uma imagem de cada indivíduo, captada em pose frontal e com uma expressão facial neutra, sendo a principal vantagem desta opção a diminuição de memória ocupada e do tempo de reconhecimento. Da simplicidade desta base de dados advém ainda a vantagem de a sua criação ser menos incómoda para os indivíduos. Como só se precisa de uma imagem de cada indivíduo, não se verifica a necessidade de expressar diferentes emoções, nem de efectuar trocas de “adereços” ou de roupa; o processo de inserção de um determinado indivíduo na base de dados fica concluído no mesmo dia.

Um sistema de reconhecimento de imagens é constituído pelas seguintes fases (ver Figura 2.1):

- Aquisição de dados através de um sistema de sensores ou de câmaras;
- Processamento de imagens (para eliminar possível ruído);
- Extracção de características mais marcantes dos dados adquiridos para reduzir a dimensionalidade dos dados, eliminando os redundantes;
- Reconhecimento de padrões, que tem por finalidade analisar as características principais, comparar representações e tomar uma decisão;

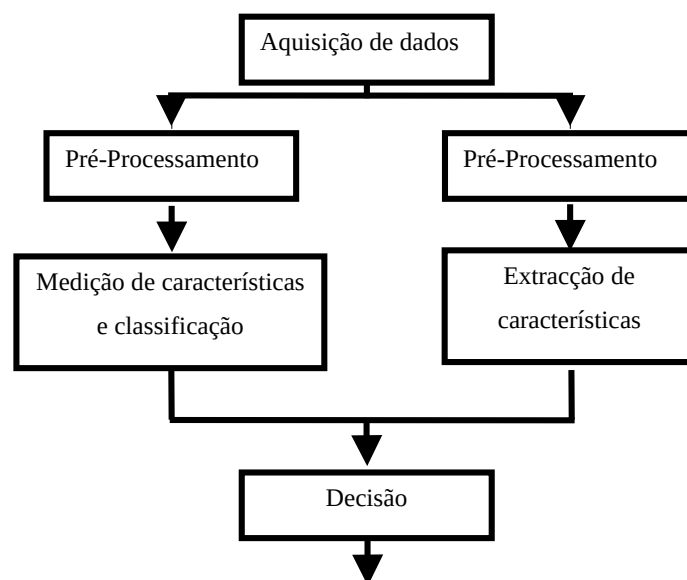


**Figura 2.1 - Esquema de Reconhecimento de Imagens.**

Após a aquisição de dados, os mesmos são separados de forma a construir a base de dados e as imagens de teste. Neste processo, as imagens que constituem a base de

dados são classificadas, formando um conjunto de treino. Este conjunto e este processo têm por finalidade afinar um classificador dotando-o da capacidade de tomar decisões relativamente às características treinadas. O conjunto de teste é definido como um conjunto de imagens (geralmente não pertencentes à base de dados) a confrontar com o conjunto de treino para ensaiar o método de reconhecimento. Nesta fase, são medidas as características consideradas fundamentais para o processo de reconhecimento em causa e, após a definição de uma medida de reconhecimento (que avalia a semelhança entre a característica testada e as características no conjunto de treino), o classificador emite uma decisão relativa ao reconhecimento.

## 2.2 Representação de Imagens

Dado o grande volume de dados envolvidos num problema de Análise e Processamento de Imagem, torna-se fulcral, para não inviabilizar o estudo, obter uma representação simples dos mesmos, por forma a reduzir a dimensão dos dados. Este objectivo pode ser alcançado através de variadas técnicas robustas e eficientes (também aplicáveis em outros domínios), tornando o problema mais simples de analisar.

A representação de imagens baseia-se em transformações lineares discretas dos dados observados.

Considere-se uma imagem  $\mathbf{x}(\mathbf{c})$  (onde  $\mathbf{c}=(i,j)$ ) com diferentes níveis de gradação da cor cinzenta constituída por  $n$  *pixels*.  $\mathbf{x}(\mathbf{c})$  denota o valor, na escala de cinzentos, do *pixel* com coordenadas  $\mathbf{c}$ . Segundo vários modelos de representação de imagens noutro espaço de dados,  $\mathbf{x}(\mathbf{c})$  é expressa como a combinação linear de  $n$  funções base  $\mathbf{a}_k$ :

$$\mathbf{x}(\mathbf{c}) = \sum_{k=1}^n \mathbf{a}_k s_k \quad (\text{Eq. 2.1})$$

onde por  $s_k$  se denotam as variáveis latentes que não são directamente observáveis, diferentes para cada imagem  $\mathbf{x}(\mathbf{c})$ , e que correspondem à solução do problema em questão.

É possível considerar uma representação matricial, conhecida por modelo estatístico linear:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (\text{Eq. 2.2})$$

bastando para tal, associar os valores de todos os *pixels* num vector coluna  $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)^T$ .

Este modelo descreve os dados observados  $\mathbf{x}$ , como sendo gerados através de um processo de mistura, representado pela matriz  $\mathbf{A}$ , que depende das variáveis independentes  $\mathbf{s}$ . Nesta representação  $\mathbf{x}$  é um vector de dimensão  $n$ ,  $\mathbf{A}$  é uma matriz rectangular de dimensão  $n \times m$  e  $\mathbf{s}$  é um vector de dimensão  $m$ .

Na grande maioria dos problemas, a matriz  $\mathbf{A}$ , designada por matriz de mistura, é desconhecida. A resolução do problema é então atingida por recurso a métodos clássicos, ou seja, é possível obter as variáveis independentes  $\mathbf{s}$  através do seguinte processo de separação:

$$\mathbf{s} = \mathbf{W}\mathbf{x} \quad (\text{Eq. 2.3})$$

onde a matriz  $\mathbf{W}$  se designa por matriz de separação.

Os métodos passíveis de utilização neste problema dividem-se, essencialmente, em *métodos de segunda ordem* e em *métodos de ordem superior*.

O facto de assumirem a gaussianaridade das variáveis, considerando apenas a informação contida nos momentos de primeira e segunda ordem dos dados observados, torna a implantação computacional dos métodos de segunda ordem mais fácil que a dos de ordem superior.

Os métodos de segunda ordem procuram encontrar uma representação dos dados utilizando apenas a informação contida na matriz de covariâncias dos dados. Assume-se que a variável  $\mathbf{x}$  segue uma distribuição gaussiana com média nula, com função de probabilidade dada pela seguinte expressão:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \quad (\text{Eq. 2.4})$$

Estes métodos permitem caracterizar a distribuição da variável através da informação contida nos momentos de primeira e de segunda ordem. Este pressuposto justifica a popularidade de técnicas como a Análise de Componentes Principais (PCA) e Análise de Factores Comuns (CFA), [Hyvärinen\_99].

Os métodos que consideram a informação contida nos momentos de ordem superior a dois são designados por métodos de ordem superior (*higher order*). Estas técnicas utilizam informação que não está contida na matriz de covariâncias dos dados. Dado que a informação que caracteriza variáveis que seguem uma distribuição normal está contida na matriz de covariâncias, estas técnicas impõem que as variáveis não sigam uma distribuição gaussiana. Redução de redundância, desconvolução cega e Análise de Componentes Independentes são alguns exemplos de métodos de ordem superior. Como métodos de ordem superior, será apenas abordada a técnica ICA, procedendo-se a uma descrição detalhada da mesma na secção 2.4.

## 2.3 Análise de Factores

A análise de factores tem por principais objectivos encontrar padrões simples nas relações entre as variáveis e reduzir o número de variáveis, por forma a explicar os dados observados (variáveis dependentes) a partir de um pequeno número de variáveis latentes com menor redundância. Esta técnica não só permite identificar relações entre os dados como também permite expressá-los de forma a evidenciar similaridades e discrepâncias entre os mesmos. Permite ainda reduzir a dimensão dos dados de tal forma que a perda de informação possa ser considerada insignificante.

A solução deste problema encontra-se determinando as variáveis latentes que contêm a informação fundamental descrita pelas variáveis independentes.

De acordo com os objectivos pretendidos, este método separa-se nos seguintes:

- Análise de factores comuns – caso se pretenda a maximização das correlações observadas;
- Análise de componentes principais – caso se pretenda a maximização da variância dos dados observados.

A utilização destas técnicas requer que os dados tenham média nula; por este motivo procede-se ao ajuste dos dados, subtraindo a média de cada dimensão aos dados. Por motivos de simplificação de escrita, nos desenvolvimentos seguintes sempre que é feita uma alusão às variáveis observadas considera-se que estas estão já centradas.

### 2.3.1 Análise de Componentes Principais

A análise de componentes principais é uma técnica que permite identificar relações entre dados, expressando-os de forma a evidenciar similaridades e diferenças entre os mesmos. Permite ainda reduzir a dimensão dos dados, de modo a que a perda de informação seja considerada insignificante.

Aplicada à análise de imagens, esta técnica considera que uma imagem com  $n$  *pixels* pode ser representada como combinação linear de  $n$  imagens-base com  $n$  *pixels* cada uma. Uma imagem base com  $n$  *pixels* é constituída por um único *pixel* activado com intensidade 1. Na combinação linear considerada, cada imagem-base tem um único *pixel* activo diferente, por forma a reconstituir a imagem original. A determinação desta combinação linear, que equivale à estimativa das componentes principais, corresponde à determinação da matriz  $\mathbf{W}$  de (Eq. 2.3). O objectivo desta técnica é então encontrar o conjunto de imagens-base que explica a maior variabilidade total de forma a que as intensidades de brilho da imagem (que correspondem aos coeficientes PCA) na nova base são não correlacionadas.

Como as fontes (variáveis independentes) são assumidas gaussianas a verosimilhança dos dados depende apenas das estatísticas ou momentos de primeira e segunda ordem, nomeadamente da matriz de covariâncias. Assim, as linhas da matriz  $\mathbf{W}$  são os autovectores (ou vectores próprios) da matriz de covariâncias dos dados,  $\Sigma$ , logo o problema de encontrar as componentes principais reduz-se à determinação dos

autovalores (ou valores próprios) de  $\Sigma$ . São os autovectores que permitem caracterizar os dados; quanto mais elevado o autovalor mais importante a informação que o autovector associado traduz relativamente aos dados.

Depois de determinados os autovalores de  $\Sigma$ , os mesmos são ordenados por ordem decrescente, organizando as componentes por ordem de significância. O autovector com autovalor associado mais elevado é considerado a primeira componente principal, e assim sucessivamente. A primeira componente principal é a combinação normalizada com variância máxima, fornecendo a máxima separação possível entre as variáveis segundo a direcção do vector próprio.

Estando as componentes organizadas por ordem de importância, é possível reduzir a dimensão dos dados descartando os autovectores com autovalores associados menores, minimizando a perda de informação. De facto, a redução dos dados é baseada na escolha heurística de um pequeno número dos primeiros vectores próprios cuja variância corresponde a uma percentagem elevada da variância total dos dados. Bartlett *et al.* [Bartlett\_01] utilizam esta técnica para proceder ao reconhecimento de imagens.

Considere-se um conjunto com  $N$  dados. Utilizando apenas os  $P$  ( $P < N$ ) autovectores com autovalores associados mais elevados, é possível reduzir a dimensão dos dados de  $N$  para  $P$ , bastando para isso utilizar a matriz constituída pelos  $P$  autovectores (ordenados segundo os autovalores associados) para escrever os dados modificados como combinação linear dos originais:

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1N}x_N \\ y_2 = a_{21}x_1 + a_{22}x_2 + \dots + a_{2N}x_N \\ \vdots \\ y_P = a_{P1}x_1 + a_{P2}x_2 + \dots + a_{PN}x_N \end{cases} \quad (\text{Eq. 2.5})$$

onde  $x_j$  ( $j=1,2,\dots,N$ ) representam os dados originais (correlacionados entre si),  $y_i$  ( $i=1,2,\dots,P$ ) representam as novas variáveis (não correlacionadas), designadas por componentes principais, e  $a_{ij}$  ( $i=1,2,\dots,P$ ,  $j=1,2,\dots,N$ ) representam os coeficientes, ou pesos, que definem cada uma das novas variáveis explicando a máxima variação dos dados.

Estas combinações lineares representam um novo sistema de coordenadas que correspondem a direcções com máxima variância (consoante o valor próprio associado).

## 2.3.2 Análise de Factores Comuns

A análise de factores comuns é uma técnica estatística que permite a determinação dos factores do problema com ruído [Pestana\_03]. De facto, na maioria dos problemas reais é impossível recolher dados sem que estes estejam afectados por erros e/ou incertezas, provocados pelo próprio investigador ou por limitações da precisão da máquina.

É então necessário proceder à reformulação do problema descrito pela (Eq. 2.2), originando:

$$\mathbf{x} = \mathbf{A}_n \mathbf{F} + \mathbf{n} \quad (\text{Eq. 2.6})$$

onde  $\mathbf{x}$  é o vector que representa os dados observados,  $\mathbf{A}_n$  é a matriz de mistura associada ao problema,  $\mathbf{F}$  os factores (variáveis independentes) e  $\mathbf{n}$  é o vector que designa a presença de ruído.

Começando por estudar as covariâncias e as correlações lineares entre as variáveis observadas, esta técnica decompõe a variância das variáveis em duas partes, descrevendo cada variável  $(x_1, x_2, \dots, x_N)$  como combinação linear dos factores comuns a todas as variáveis  $(F_1, F_2, \dots, F_M)$  e de um factor único a cada variável  $(U_1, U_2, \dots, U_N)$ :

$$\begin{cases} x_1 = b_{11}F_1 + b_{12}F_2 + \dots + b_{1M}F_M + U_1 \\ x_2 = b_{21}F_1 + b_{22}F_2 + \dots + b_{2M}F_M + U_2 \\ \vdots \\ x_N = b_{N1}F_1 + b_{N2}F_2 + \dots + b_{NM}F_M + U_N \end{cases} \quad (\text{Eq. 2.7})$$

onde  $b_{ij}$  ( $i = 1, 2, \dots, N$ ,  $j = 1, 2, \dots, M$ ) representam as correlações entre os factores comuns e as variáveis observadas e  $U_j$  ( $j = 1, 2, \dots, N$ ) representa a parte da variável que não é explicada pelos factores comuns.

A variância da variável que é explicada pelos factores comuns designa-se por comunalidade da variável e o valor dos coeficientes indica a quantidade de variância de uma variável que é partilhada, pelo menos, com outra variável do conjunto, [Pestana\_03]. Os factores únicos explicam o ruído associado ao problema, atribuindo uma fracção de ruído total a cada variável do modelo.

## 2.4 Análise de Componentes Independentes

Tendo uma base em comum com a Análise de Componentes Principais e com a Análise de Factores Comuns, a técnica de Análise de Componentes Independentes é mais robusta e firme que as anteriores pois, tendo em consideração a informação dos dados contida nos momentos de ordem superior a dois, permite a resolução de problemas que as abordagens referidas não possibilitam.

Sendo comparativamente mais recente é passível de utilização nas mais diversas áreas científicas e técnicas, tais como Processamento de Sinal e de Imagem, Problemas de Séries Temporais e Problemas Financeiros, Separação de Fontes (telecomunicações), Extração de Características, entre outros, como referido em [Hyvärinen\_01].

São exemplos de resolução através desta técnica misturas de sinais de fala simultâneos captados através de vários microfones, ondas cerebrais registadas por múltiplos sensores, séries temporais obtidas num processo industrial ou ainda a identificação e/ou reconhecimento de imagens presentes numa base de dados.

Esta técnica estatística e computacional baseia-se na construção de um modelo explicativo dos dados multivariados observados, assumindo a independência estatística e a não-gaussianidade das componentes que o definem, designadas por fontes ou componentes independentes. As variáveis em estudo são descritas como uma combinação linear ou processo de mistura desconhecido das componentes independentes (e por consequência não correlacionadas) que o método ICA propõe encontrar.

Este modelo descreve como os dados são gerados a partir das componentes independentes  $s_i$  ( $i = 1, 2, \dots, n$ ). Para as encontrar, as imagens (dados)  $x$  são representadas como uma combinação linear de imagens-base estatisticamente independentes,  $s_i$  ( $i = 1, 2, \dots, n$ ), descrita pela (Eq. 2.1).

Aplicada ao processamento de imagens, nomeadamente ao reconhecimento de faces humanas, a finalidade desta técnica é decompor uma imagem como combinação de imagens-base estatisticamente independentes (ou tanto quanto possível),  $s$ .

Como estas variáveis latentes não podem ser observadas directamente e dado que se dispõe apenas de informação relativa aos dados observados  $\mathbf{x}$ , a matriz de mistura  $\mathbf{A}$ , que transforma as componentes independentes  $\mathbf{s}$  na imagem original  $\mathbf{x}$ , é desconhecida, o que força a necessidade de estimar as componentes independentes. Caso a matriz  $\mathbf{A}$  fosse conhecida e invertível, era trivial encontrar as componentes independentes, recorrendo à equação (Eq. 2.3), bastando considerar  $\mathbf{W} = \mathbf{A}^{-1}$ .

A Análise de Componentes Independentes propõe então um modelo para estimar  $\mathbf{A}$  e  $\mathbf{s}$ . Na prática, após estimar a matriz  $\mathbf{A}$  é possível calcular a sua inversa  $\mathbf{W}$  e encontrar as estimativas das componentes independentes através do processo de separação descrito pela (Eq. 2.3) e como se verifica na Figura 2.2:

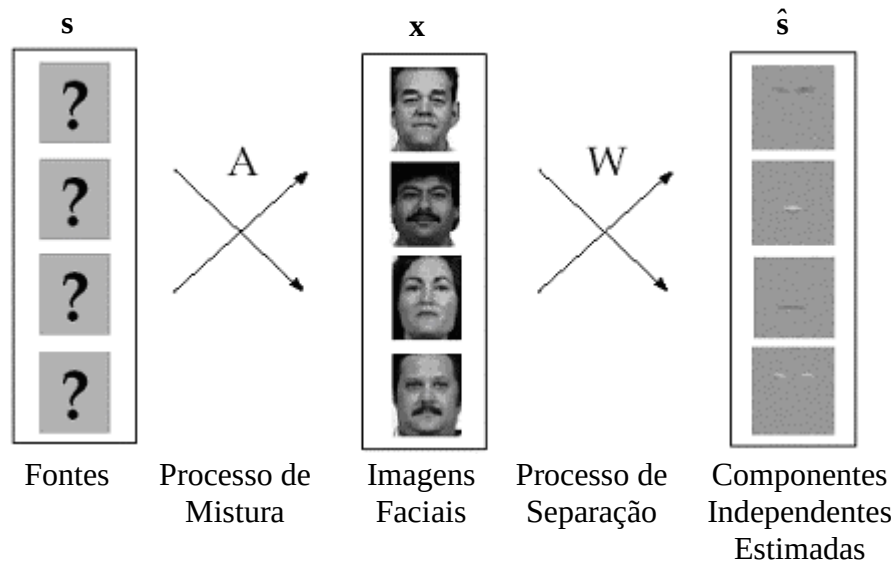


**Figura 2.2 - Modelo de Estimação das Componentes Independentes. Adaptado de [Bartlett\_98].**

Assim, os algoritmos sugeridos para esta técnica têm por finalidade concretizar um processo de estimação das componentes independentes. Este processo depende da estimação da matriz  $\mathbf{W}$ , que representa o processo de separação, por forma a obter-se:

$$\hat{\mathbf{s}} = \mathbf{W}\mathbf{x} \quad (\text{Eq. 2.8})$$

onde  $\hat{\mathbf{s}}$  denota uma estimativa das componentes independentes  $\mathbf{s}$ .

A Figura 2.3 ilustra este processo de estimação:

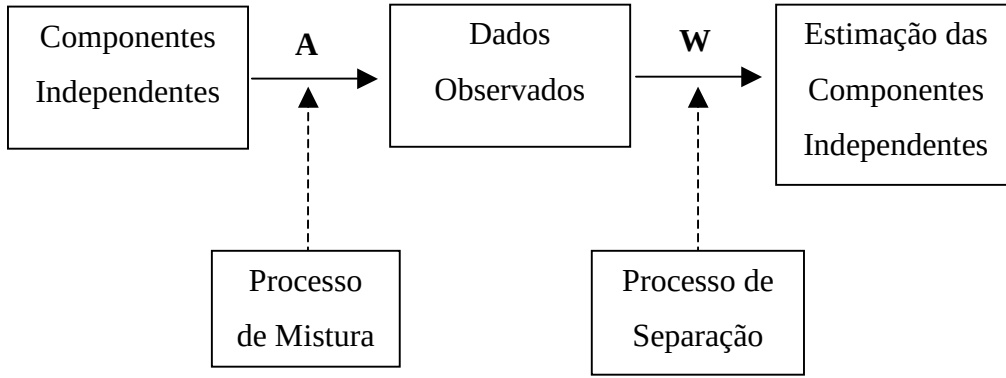


**Figura 2.3 - Modelo de separação das componentes independentes. Adaptado de [Bartlett\_98].**

Após a estimação das componentes independentes, cada face (imagem) contida no conjunto de dados  $\mathbf{x}$  é representada como combinação linear com coeficientes  $\mathbf{b}_i$  ( $i = 1, 2, \dots, n$ ) das imagens base  $\hat{\mathbf{s}}_i$  ( $i = 1, 2, \dots, n$ ) estimadas:

$$\text{Imagem de Face} = b_1 * \hat{s}_1 + b_2 * \hat{s}_2 + \dots + b_n * \hat{s}_n$$

Representação ICA:  $B = (b_1, b_2, \dots, b_n)$

**Figura 2.4 - Representação das faces como combinação linear das componentes independentes. Adaptado de [Bartlett\_98].**

Deste modo, a técnica ICA consiste em encontrar a matriz de pesos  $\mathbf{W}$  de forma a que as componentes  $\mathbf{s}_i$  ( $i = 1, 2, \dots, n$ ) sejam tão independentes entre si quanto possível. Esta independência é atingida através da maximização de uma função que mede a independência estatística, como será abordado na secção 2.4.1.

Considerem-se as variáveis aleatórias observadas e centradas  $x_1, x_2, \dots, x_n$ . Estas variáveis são estatisticamente independentes se a sua função de densidade conjunta se puder escrever como produto das funções de densidade marginais de cada variável:

$$f(x_1, x_2, \dots, x_n) = f_1(x_1)f_2(x_2)\dots f_n(x_n) \quad (\text{Eq. 2.9})$$

onde  $f_i(x_i)$ ,  $i = 1, 2, \dots, n$  designa a função de densidade marginal da variável  $x_i$ ,  $i = 1, 2, \dots, n$ .

Contudo, geralmente, é impossível encontrar uma transformação  $W$  que conduza a uma independência estatística total. Por este motivo, para avaliar a independência das componentes, recorre-se a funções de medida da independência.

Segundo, Hyvärinen, [Hyvärinen\_01], a estimativa  $\hat{s}$  das componentes independentes  $s$ , segue dois princípios fundamentais:

- As estimativas das componentes  $\hat{s}$  devem ser tão independentes quanto possível.
- As estimativas das componentes independentes  $\hat{s}$  devem ser não gaussianas (com a possível excepção de apenas uma).

De acordo com o Teorema do Limite Central, a soma de diversas variáveis aleatórias independentes geralmente possui uma distribuição mais próxima da distribuição normal do que cada uma das variáveis. Assim, esta técnica procura, iterativamente, extrair variáveis aleatórias a partir dos dados observados de forma a que sejam não gaussianas (tanto quanto possível). Tendo em conta este teorema, os dois princípios fundamentais da técnica ICA podem então ser satisfeitos através da utilização de uma função objectivo que mede a não gaussianaridade das componentes (e por consequência avalia a independência estatística) e através de um algoritmo iterativo de optimização.

A escolha da função objectivo e do algoritmo de optimização é independente, dado que é possível utilizar qualquer um dos métodos clássicos de optimização com uma função objectivo e cada um desses métodos pode ser utilizado com várias funções objectivo. Deste modo, a escolha da função objectivo e do respectivo algoritmo de optimização é condicionada por:

- *Métodos de estimação das componentes independentes*: estimar todas as componentes em paralelo ou estimar uma-a-uma.
- *Propriedades estatísticas da função objectivo*: consistência, variância assintótica e robustez.
- *Propriedades dos algoritmos de optimização*: eficiência, memória utilizada, rapidez de convergência e estabilidade numérica do algoritmo.

## 2.4.1 Métodos de Estimação

A escolha da função objectivo a utilizar na estimação das componentes independentes depende dos métodos de estimação adoptados.

Nesta secção são apresentados alguns métodos de estimação. Dado que é possível encontrar a descrição destes métodos em vários livros de estatística, apresentam-se apenas descrições sumárias que os permitam distinguir e identificar as possíveis funções objectivo utilizadas para medir a não gaussianidade. Assim, são apresentados o método da Minimização da Informação Mútua e o método de Maximização de Não Gaussianidade.

### 2.4.1.1 Método de Minimização da Informação Mútua

O método da minimização da informação mútua baseia os seus princípios na teoria da informação. A teoria da informação centra o seu estudo no armazenamento e transmissão de informação. Considere-se um sistema de comunicação com  $m_1, m_2, \dots, m_L$  como possíveis mensagens, com probabilidades de ocorrência  $P_1, P_2, \dots, P_L$ . Neste sistema, como é obvio,  $P_1 + P_2 + \dots + P_L = 1$ . Quando a mensagem de ordem  $k$  é transmitida, define-se como quantidade de informação transportada,  $I_k$ , como:

$$I_k = -\log_2(P_k) \quad (\text{bits}) \quad (\text{Eq. 2.10})$$

A informação média, ou entropia, associada à transmissão de cada uma das  $L$  mensagens do sistema de comunicação é a média ponderada da informação de cada mensagem, ([Taub\_71] – capítulo 13):

$$H = -\sum_{j=1}^L P_j \log_2(P_j) \quad (\text{bits/mensagem}) \quad (\text{Eq. 2.11})$$

Considere-se agora um sistema de comunicação contínuo. Neste caso, a quantidade média de informação de um vector aleatório contínuo  $\mathbf{x}$  (de dimensão  $N$ ),

com função de densidade de probabilidade  $f_x(\xi)$  é medida através da função de entropia diferencial, denotada por  $H$  e definida por:

$$H(\mathbf{x}) = - \int f_x(\xi) \log_2(f_x(\xi)) d\xi \quad (\text{Eq. 2.12})$$

A entropia diferencial atinge o seu valor máximo se e só se  $f_x(\xi)$  for a função de densidade de probabilidade gaussiana. Ou seja, se  $\mathbf{x}$  e  $\mathbf{y}$  forem vectores aleatórios contínuos com a mesma variância e se  $\mathbf{x}$  for gaussiano, então  $H(\mathbf{x}) \geq H(\mathbf{y})$ .

A entropia diferencial do processo  $\mathbf{y} = \mathbf{W}\mathbf{x}$ , sendo  $\mathbf{W}$  uma transformação não singular, pode ser definida em função da entropia diferencial de  $\mathbf{x}$ , como:

$$H(\mathbf{y}) = H(\mathbf{x}) + \log_2(\det \mathbf{W}) \quad (\text{Eq. 2.13})$$

Baseada no conceito de entropia surge a neguentropia, uma medida que assume valor nulo para variáveis gaussianas e valores positivos para variáveis não gaussianas.

A neguentropia, denotada por  $J$ , de um vector aleatório contínuo  $\mathbf{x}$  é definida por:

$$J(\mathbf{x}) = H(\mathbf{x}_G) - H(\mathbf{x}) \quad (\text{Eq. 2.14})$$

onde  $\mathbf{x}_G$  é um vector aleatório contínuo gaussiano que partilha os dois primeiros cumulantes (média e matriz de covariâncias) com  $\mathbf{x}$ , e  $H(\cdot)$  é a entropia diferencial do vector. A neguentropia é uma medida invariante a transformações lineares invertíveis:

$$J(\mathbf{y}) = J(\mathbf{W}\mathbf{x}) = J(\mathbf{x}) \quad (\text{Eq. 2.15})$$

É com fundamento nestes conceitos e no conceito de entropia conjunta que Shannon [Shannon\_48] introduziu o conceito de “proporção de transmissão de informação”, comumente conhecido como informação mútua. A informação mútua entre duas variáveis aleatórias pode, qualitativamente, representar uma medida de quão bem uma variável explica a outra. Ou seja, é uma medida da informação que uma variável pertencente a um conjunto de variáveis aleatórias contém em relação a outra variável do mesmo conjunto:

$$I(\mathbf{x}) = \int f_{\mathbf{x}}(\xi) \log_2 \left( \frac{f_{\mathbf{x}}(\xi)}{\prod_{i=1}^n f_{\mathbf{x}_i}(\xi_i)} \right) d\xi = \sum_{i=1}^n H(\mathbf{x}_i) - H(\mathbf{x}) \quad (\text{Eq. 2.16})$$

Recorrendo à definição de neguentropia e de entropia de uma variável gaussiana, Comon [Comon\_94] representa a informação mútua em função das neguentropias (marginais e conjunta) e da matriz de covariâncias:

$$I(\mathbf{x}) = J(\mathbf{x}) - \sum_{i=1}^n J(\mathbf{x}_i) + \frac{1}{2} \log_2 \left( \frac{\prod_{i=1}^n \mathbf{V}_{ii}}{\det \mathbf{V}} \right) \quad (\text{Eq. 2.17})$$

onde  $\mathbf{V}$  representa a matriz de covariâncias do vector aleatório contínuo  $\mathbf{x}$ .

A prova da igualdade entre as duas últimas expressões pode ser encontrada em [Comon\_94].

Da última expressão resulta que se  $\mathbf{x}$  seguir uma distribuição gaussiana então:

$$I(\mathbf{x}) = \frac{1}{2} \log_2 \left( \frac{\prod_{i=1}^n \mathbf{V}_{ii}}{\det \mathbf{V}} \right) \quad (\text{Eq. 2.18})$$

Neste caso a informação mútua é nula se e só se  $\det \mathbf{V} = \prod_{i=1}^n \mathbf{V}_{ii}$ . Ou seja, a independência estatística atinge-se obrigando a matriz  $\mathbf{V}$  a ser diagonal.

Recorde-se a equação  $\mathbf{s} = \mathbf{W}\mathbf{x}$ . Então, através da propriedade da neguentropia ser invariante para transformações lineares invertíveis (Eq. 2.15) e atendendo à definição (Eq. 2.16) a informação mútua de  $\mathbf{s}$  pode ser representada como:

$$I(\mathbf{s}) = \sum_{i=1}^n H(\mathbf{s}_i) - H(\mathbf{x}) - \log_2(\det \mathbf{W}) \quad (\text{Eq. 2.19})$$

Daqui resulta que  $I(\mathbf{s}) \leq I(\mathbf{x})$ .

Se considerarmos que  $\mathbf{x}$  é um vector aleatório esférico (com matriz de covariâncias igual à identidade) que segue uma distribuição não gaussiana, pela (Eq. 2.15) tem-se que:

$$\sum_{i=1}^n J(\mathbf{s}_i) \geq \sum_{i=1}^n J(\mathbf{x}_i) \quad (\text{Eq. 2.20})$$

Desta forma, o método de minimização da informação mútua permite obter a independência estatística das componentes de  $\mathbf{x}$  (  $\mathbf{s}$  ) maximizando as somas das neguentropias marginais.

#### **2.4.1.2      *Método de Maximização da Não Gaussianidade***

Recorrendo a este método as componentes independentes são estimadas pela determinação de direcções nas quais se verificam os valores máximos de não gaussianidade encontrados através dos algoritmos de optimização (abordados na secção 2.4.2). A *kurtosis* e a neguentropia são duas medidas que permitem medir a não gaussianidade das componentes independentes.

A *kurtosis* (cumulante de 4ª ordem), denotada por *kurt*, de um vector aleatório  $\mathbf{x}$  é dada por:

$$kurt(\mathbf{x}) = E(\mathbf{x}^4) - 3[E(\mathbf{x}^2)]^2 \quad (\text{Eq. 2.21})$$

A *kurtosis*, por vezes designada por coeficiente de excesso, é uma medida que permite avaliar o grau de achatamento da função de distribuição de uma variável por comparação com a distribuição normal padrão (ou *standard*). Um vector aleatório  $\mathbf{x}$  com distribuição normal *standard* apresenta um valor de *kurtosis* nulo. Deste modo, quando a *kurtosis* é negativa o grau de achatamento é maior do que o da distribuição normal, sendo que quando o grau de achatamento é mais agudo que o da distribuição normal a *kurtosis* é positiva. Assim, em geral, considera-se o valor absoluto da *kurtosis* para medir a não gaussianidade de uma variável.

Contudo, dado ser definida a partir dos momentos da variável, esta medida é muito sensível a *outliers*, sendo bastante afectada por valores extremos e dados mal medidos ou introduzidos erroneamente. Associadas a esta limitação surgem outras medidas de não gaussianidade, como, por exemplo, a neguentropia, dada pela (Eq. 2.14).

Como já foi referido, a neguentropia é uma medida que assume valor nulo para variáveis gaussianas e valores positivos para variáveis não gaussianas. Pela sua definição, a neguentropia é passível de ser utilizada como medida de não gaussianidade. Contudo, apesar de ser um estimador robusto da não gaussianidade de uma variável, o cálculo do integral da entropia diferencial (ver (Eq. 2.12)) aumenta o esforço computacional. Não obstante, dada a pequena sensibilidade a *outliers* é preferível recorrer, em detrimento do uso da *kurtosis*, a aproximações da neguentropia, definindo novas medidas que apesar de serem menos robustas que a neguentropia oferecem maior confiança que a *kurtosis*. Hyvärinen, [Hyvärinen\_99], sugere as aproximações da neguentropia através de cumulantes e através de funções não polinomiais. A aproximação da neguentropia através de cumulantes é dada pela seguinte expressão:

$$J(\hat{s}) \approx \frac{1}{2} \left[ E(\hat{s}^3) \right]^2 + \frac{1}{48} \left[ kurt(\hat{s}) \right]^2 \quad (\text{Eq. 2.22})$$

onde  $\hat{s}$  representa as estimativas dos coeficientes ICA.

Através de funções não polinomiais, é sugerida a seguinte aproximação à neguentropia:

$$J(\hat{s}) \approx \sum_{i=1}^p k_i \left\{ E[G_i(\hat{s})] - E[G_i(v)] \right\}^2 \quad (\text{Eq. 2.23})$$

onde,  $k_i$  são constante positivas,  $\hat{s}$  representa as estimativas dos coeficientes ICA e  $v$  é uma variável gaussiana padrão.

Utilizando apenas uma função de ordem superior à quadrática, tem-se:

$$J(\hat{s}) \propto \left\{ E[G(\hat{s})] - E[G(v)] \right\}^2 \quad (\text{Eq. 2.24})$$

A obtenção de um estimador robusto da não gaussianidade depende agora da escolha da função não polinomial  $G$ . Baseando-se na robustez e na variância assintótica do estimador, o autor referido propõe as seguintes funções:

$$G(z) = \log \cosh(az), \quad \text{com } 1 < a < 2 \quad (\text{Eq. 2.25})$$

e

$$G(z) = \exp\left(-\frac{z^2}{2}\right) \quad (\text{Eq. 2.26})$$

## 2.4.2 Algoritmos de Optimização

Como já foi referido a escolha do algoritmo de optimização é condicionada pelas suas propriedades.

O algoritmo do gradiente e o algoritmo do ponto fixo (*fastICA*) são os mais comuns na determinação da estimativa das componentes independentes sendo que a adaptação de cada um ao problema em questão deve ser ajustada ao método de estimação utilizado.

A descrição dos algoritmos de optimização considera-se fora do âmbito deste trabalho, no entanto segue-se uma comparação que tem por finalidade permitir escolher um dos dois para aplicação neste trabalho.

O algoritmo *fastICA* é um método computacional altamente eficiente para atingir a estimação de componentes independentes. Usa um esquema de repetição do ponto-fixa desenvolvido em experiências independentes com a finalidade de ser mais rápido que métodos de gradiente convencionais. No que respeita à convergência, o algoritmo *fastICA* apresenta uma convergência aproximadamente quadrática enquanto que o algoritmo do gradiente tem uma convergência linear. É ainda uma vantagem do primeiro o facto de determinar as estimativas das componentes independentes utilizando qualquer função objectivo não linear, contrariamente ao outro algoritmo referido. Além destas desvantagens, o algoritmo do gradiente necessita da definição do tamanho do passo iterativo para calcular determinados parâmetros, o que não acontece no algoritmo *fastICA*.

### 2.4.3 Representação Matricial dos Dados

Cada imagem corresponde a uma matriz, onde cada elemento escalar representa uma intensidade na escala de níveis cinzentos. Contudo, como a técnica ICA actua sobre o conjunto de dados, ou seja sobre todas as imagens, é necessário escolher outra representação matricial para os mesmos. Considera-se então que cada imagem é representada por um vector com dimensão igual ao número de *pixels* que a definem.

Considere-se uma imagem de dimensão  $N$  por  $M$  :

$$\text{Imagem} = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,M} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N,1} & x_{N,2} & \cdots & x_{N,M} \end{pmatrix} \quad (\text{Eq. 2.27})$$

onde  $x_{i,j}$ ,  $i = 1, 2, \dots, N$  e  $j = 1, 2, \dots, M$ , representa a intensidade do *pixel*  $(i,j)$ . Se for uma imagem no sistema RGB,  $x_{i,j}$  é um vector 3-dimensional, caso seja uma imagem em escala de cinzentos,  $x_{i,j}$  representa apenas um valor escalar.

Desta forma, uma imagem de dimensão  $N$  por  $M$  pode ser representada pela concatenação das suas linhas resultando num vector  $NM$  dimensional:

$$\text{Vector Imagem} = (x_{1,1}, \dots, x_{1,M}, x_{2,1}, \dots, x_{2,M}, \dots, x_{N,1}, \dots, x_{N,M}) \quad (\text{Eq. 2.28})$$

onde as linhas das imagens são colocadas uma após outra de forma a construir uma imagem unidimensional (os valores deste vector são os valores de intensidade da imagem original).

Segundo Bartlett *et al.*, [Bartlett\_01] e [Bartlett\_98], a forma de dispor estes vectores na matriz de dados  $\mathbf{X}$ , sobre a qual a técnica ICA actua, conduz às arquitecturas I e II, que se descrevem a seguir. A principal diferença entre estas reside no facto de na primeira se verificar independência entre as imagens e na segunda existir independência entre os *pixels* das imagens.

### 2.4.3.1 *Arquitetura I*

Segundo esta arquitetura cada vector linha da matriz de dados  $\mathbf{X}$  corresponde a uma imagem da base de dados, como se ilustra na Figura 2.5:

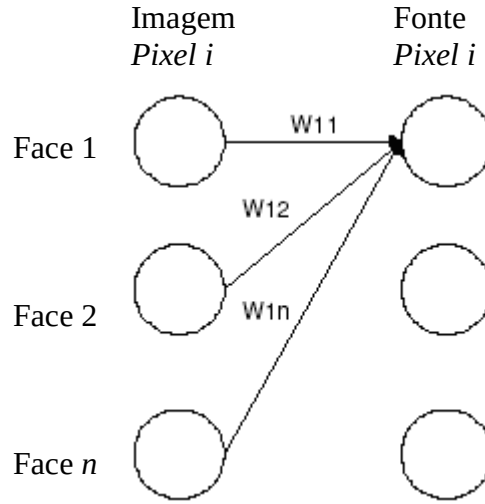


**Figura 2.5 - Separação das fontes da imagem obtendo as componentes independentes nas linhas de  $\hat{\mathbf{S}}$ . Adaptado de [Bartlett\_98].**

Considere-se um conjunto de  $P$  imagens, a matriz  $\mathbf{X}$  assume a seguinte forma:

$$\mathbf{X} = \begin{pmatrix} \text{Vector Imagem 1} \\ \text{Vector Imagem 2} \\ \vdots \\ \text{Vector Imagem P} \end{pmatrix} = \begin{pmatrix} x_1^1 & x_2^1 & \cdots & x_{N,M}^1 \\ x_1^2 & x_2^2 & \cdots & x_{N,M}^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^P & x_2^P & \cdots & x_{N,M}^P \end{pmatrix} \quad (\text{Eq. 2.29})$$

onde  $x_i^p$  representa a intensidade do índice  $i$  ( $i=1,2,\dots,NM$ ) da imagem  $p$  ( $p=1,2,\dots,P$ ) na representação vectorial descrita na (Eq. 2.28).

Aplicando a técnica ICA aos dados estruturados de acordo com esta arquitetura, a matriz de separação  $\mathbf{W}$  encontra-se impondo a independência entre as linhas de  $\hat{\mathbf{S}} = \mathbf{W}\mathbf{X}$ . Assim, as estimativas encontradas das imagens-base que representam as faces são independentes e correspondem às linhas da matriz  $\mathbf{W}$ .

### 2.4.3.2 *Arquitectura II*

Nesta arquitectura cada vector coluna da matriz dos dados  $\mathbf{X}$  corresponde a uma imagem da base de dados, como se exemplifica na Figura 2.6:

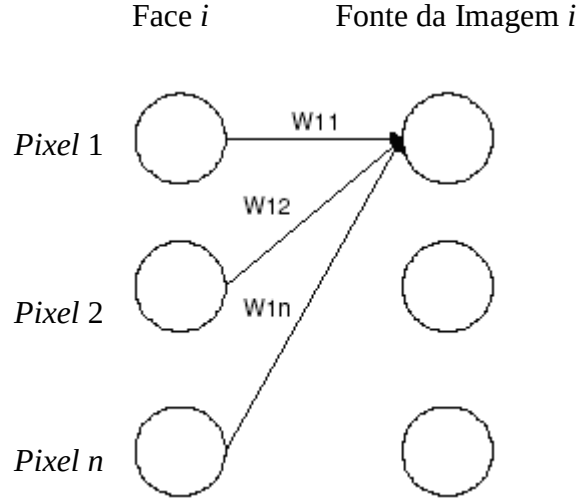


**Figura 2.6 - Separação das fontes dos *pixels* obtendo factorial *face code* nas colunas de  $\hat{\mathbf{S}}$ . Adaptado de [Bartlett\_98].**

Considerando-se novamente um conjunto de  $P$  imagens, a matriz  $\mathbf{X}$  assume a seguinte forma:

$$\mathbf{X} = \begin{pmatrix} \text{Vector Pixel 1} \\ \text{Vector Pixel 2} \\ \vdots \\ \text{Vector Pixel NM} \end{pmatrix} = \begin{pmatrix} x_1^1 & x_1^2 & \cdots & x_1^P \\ x_2^1 & x_2^2 & \cdots & x_2^P \\ \vdots & \vdots & \ddots & \vdots \\ x_{N.M}^1 & x_{N.M}^2 & \cdots & x_{N.M}^P \end{pmatrix} \quad (\text{Eq. 2.30})$$

onde  $x_i^p$  representa a intensidade do índice  $i$  ( $i=1,2,\dots,NM$ ) da imagem  $p$  ( $p=1,2,\dots,P$ ) na representação vectorial descrita na (Eq. 2.28). Nesta representação são construídos  $NM$  vectores de dimensão  $P$ , que consistem na intensidade do mesmo *pixel* em cada uma das  $P$  imagens, obtendo assim um vector para cada *pixel* onde cada elemento corresponde a uma imagem diferente.

Aplicando a técnica ICA a esta estrutura dos dados, as componentes independentes estimadas correspondem às colunas da matriz  $\mathbf{W}$ .

### 2.4.3.3 *Comparação das Architecturas*

Tendo por critério a memória utilizada e o tempo de processamento, a escolha da arquitectura utilizada para estruturar os dados depende da dimensão das matrizes calculadas durante o processo ICA.

Considerando-se que a base de dados é constituída por  $n$  imagens, cada uma com  $m$  pixels, a escolha da arquitectura recai sobre a arquitectura I no caso de  $n < m$ , caso contrário a arquitectura II é vantajosa.

A justificação desta escolha prende-se com os seguintes factos:

- Na arquitectura I, a matriz  $\mathbf{X}$  tem dimensão  $n \times m$ , logo as matrizes geradas durante a aplicação da técnica ICA têm dimensão  $n^2$ .
- Na arquitectura II, a matriz  $\mathbf{X}$  tem dimensão  $m \times n$ , logo as matrizes geradas durante a aplicação da técnica ICA têm dimensão  $m^2$ .

Considere-se a Arquitectura II para representar os dados. Após redução da dimensionalidade destes (descartando os autovectores menos significantes) verifica-se perda de informação na reconstrução dos dados, o que não acontece com a arquitectura I. Por este motivo a arquitectura II é geralmente utilizada para compressão de dados e a primeira para reconhecimento de imagens.

# 3 SÍNTESE DE VISTAS FRONTAIS

## 3.1 Revisão do Estado da Arte

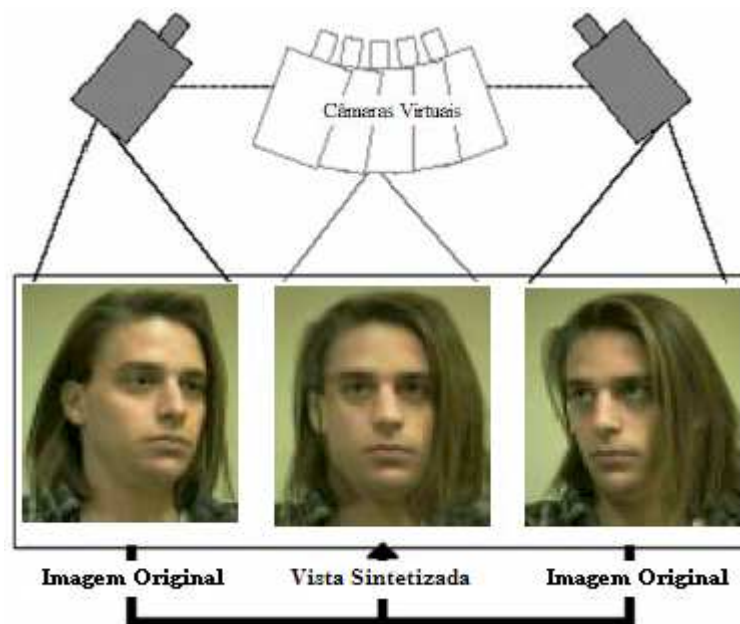


Figura 3.1 - Exemplo de Síntese de Vistas Intermédias, adaptado de [Seitz\_97].

Uma vez que as imagens decimadas da sequência vídeo podem ser não frontais, defrontamo-nos com a necessidade de criar vistas frontais para que possam ser comparadas com as imagens da base de dados.

Vários autores sugerem a construção de modelos 3D permitindo assim gerar uma imagem de um indivíduo de qualquer ponto de vista. Contudo, esta metodologia não será alvo de atenção neste trabalho, pois não é objectivo principal do mesmo desenvolver nem aperfeiçoar técnicas de síntese de imagens, mas sim implementá-las.

Ullman e Basri, [Ullman\_91], demonstram que novas vistas podem ser expressas como combinação linear entre outras vistas da mesma cena. Contudo a técnica descrita requer correspondência completa entre as imagens originais para efectuar a síntese da nova vista. Esta limitação inviabiliza a metodologia sugerida dado que na maioria das vezes é impossível estabelecer correspondência completa.

Chen e Williams, [Chen\_91], propõem a síntese de vistas com base na interpolação linear de pontos de correspondência entre as imagens originais. Porém, visto assentar na interpolação linear, a vista frontal sintetizada através desta técnica não produz uma nova vista do mesmo objecto. Esta limitação é abordada mais à frente na secção 3.2.

Lam e Yan, [Lam\_98], aplicam o modelo *snake* com o objectivo de extrair os contornos faciais das imagens originais, posteriormente detectam 15 pontos característicos na face (entre eles, os cantos dos lábios e os cantos dos olhos) através dos quais se constrói um modelo 3D. Uma imagem facial não frontal é transformada numa vista frontal por recurso ao modelo construído. A detecção de 15 pontos característicos e o facto de só poder ser aplicada a imagens quase frontais são duas das limitações desta técnica.

Beymer *et al.*, [Beymer\_93] e [Beymer\_93a], sugerem a construção de uma vista virtual baseada na teoria de que qualquer vista 2D pode ser expressa como combinação linear de outras vistas conhecidas. Esta metodologia apresenta como limitações o facto de requerer mais do que 3 vistas para a construção de uma nova imagem e de ser necessário especificar um largo número de correspondências entre as diferentes imagens. A tarefa de estabelecer um grande número de pontos de correspondência entre imagens faciais é, por si só, uma tarefa difícil e morosa. É necessário ter o cuidado de escolher devidamente os pontos, para que estes não recaiam sobre uma zona uniforme, culminando, na maioria dos casos, na escolha de pontos característicos, tais como os cantos dos lábios, dos olhos e das sobrancelhas, entre outros.

Para contornar as dificuldades inerentes ao estabelecimento de pontos de correspondência, Feng e Yuen, [Feng\_00], apresentam um algoritmo para detectar pontos-chave (*landmarks*) faciais, os quais são utilizados para estimar a orientação da face numa imagem facial. Após determinada uma aproximação da posição da face, os autores propõem a construção de um modelo 3D através do qual é possível transformar a imagem inicial numa pose frontal.

O método proposto por estes autores ultrapassa as limitações do anterior no sentido em que necessita apenas de uma imagem para construir a face frontal e dado que a mesma é obtida por recurso a um modelo 3D. Além destas vantagens, os autores defendem também como vantagem o facto de só serem necessários 3 pontos-chave faciais, não sendo necessário estabelecer correspondências.

Seitz e Dyer, [Seitz\_96] e [Seitz\_95], recorrem a técnicas de *morphing* baseadas na interpolação geométrica das imagens para, através de duas imagens de poses diferentes do mesmo objecto e a partir de uma correspondência entre *pixels* das duas imagens, criar novas vistas intermédias (ver Figura 3.1). No pressuposto do objecto ser simétrico, esta técnica pode ser utilizada mesmo quando se possui apenas uma vista não frontal do objecto, bastando para tal criar a imagem simétrica lateralmente da que se possui, sendo assim possível gerar qualquer vista intermédia entre as duas, incluindo a vista frontal. A utilização de apenas uma imagem original na síntese de uma vista frontal conduz a duas vantagens: não haver necessidade de estimar a pose lateral da face e não ser preciso determinar o valor de um parâmetro  $s$ , que indica a distância da imagem sintetizada em relação às imagens originais (este parâmetro que é introduzido na secção 3.2.1, assume, neste caso, o valor 0.5).

Para evitar o recurso, desnecessário, a modelos 3D e como os resultados apresentados por Seitz e Dyer são válidos para o presente trabalho, foi seguida a metodologia *View-Morphing* apresentada por estes.

## 3.2 View-Morphing

Define-se *Image Morphing* como uma técnica que permite sintetizar transições suaves entre duas imagens. Segundo esta técnica, a síntese de uma imagem intermédia é conseguida através da deformação das imagens originais e do cálculo da média dos valores relativos à intensidade dos *pixels* na imagem criada.

A Figura 3.2 apresenta a síntese de uma imagem frontal, entre uma imagem e a sua simétrica, através desta técnica.

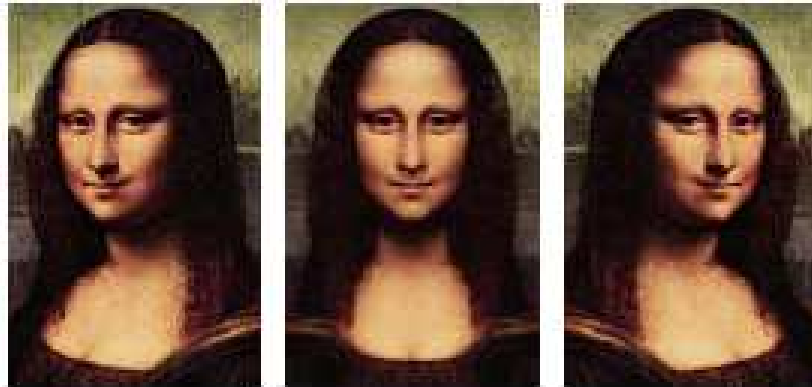


**Figura 3.2 - Síntese de imagem através da técnica *Image Morphing* entre uma vista de Mona Lisa e a sua reflexão, provocando contracção horizontal e vertical da face e do busto, [Seitz\_97].**

Embora esta técnica revele a capacidade de gerar imagens visualmente convincentes, verifica-se que após interpolação linear a duas dimensões os mapas de projecção (ver secção 3.2.1) não são preservados, o que provoca um efeito de curvatura das rectas nas imagens intermédias, como se ilustra na Figura 3.3:

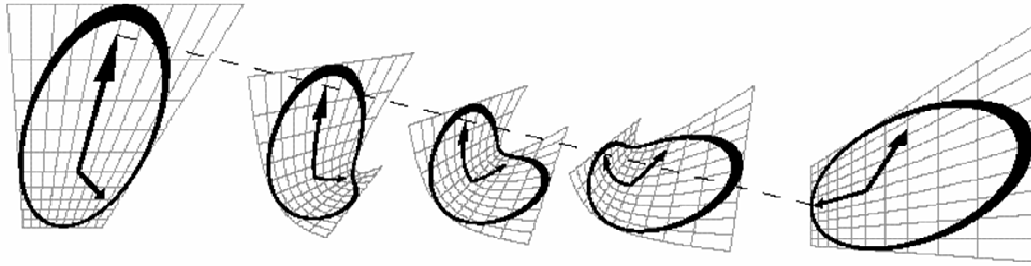


**Figura 3.3 - Distorção provocada na síntese de imagens através da técnica *Image Morphing*. Ilustra-se uma interpolação linear entre duas vistas de um relógio (imagem mais à esquerda e mais à direita) que provoca um efeito de curvatura nas imagens mais intermédias. A linha a tracejado representa o percurso de uma característica em diversas imagens intermédias, [Seitz\_97].**

De facto, uma das falhas cruciais com que esta metodologia se debate deve-se ao facto de não assegurar que alterações relativas à pose do objecto e ao ponto de vista sejam consideradas nas imagens sintetizadas, o que conduz a que possíveis transições entre as imagens originais não pareçam realistas na vista intermédia.

Estas limitações, inerentes ao *Image Morphing*, levam a admitir que o *morph* criado entre duas vistas diferentes do mesmo objecto não produz uma nova vista do mesmo objecto, como se ilustra na Figura 3.2 onde a imagem sintetizada por esta técnica apresenta uma contracção horizontal e vertical da face e do busto.

Esta falha despoletou estudos e investigações, [Seitz\_96], [Seitz\_95] e [Seitz\_97] que conduziram ao conceito de *View Morphing*.

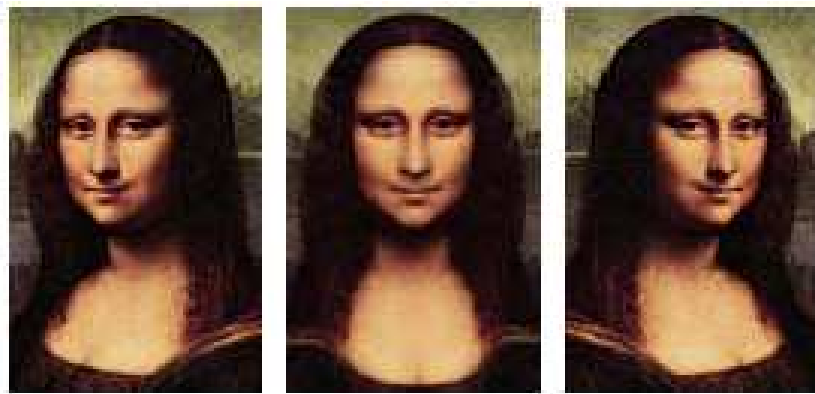


**Figura 3.4 - Síntese de imagem através da técnica *View Morphing*, entre uma vista da Mona Lisa e a sua reflexão, produzindo uma vista mais realística, [Seitz\_97].**

A metodologia *View Morphing* assenta numa interpolação baseada no *Image Morphing*, a diferença entre elas recai no facto de a primeira ser uma transformação que preserva a forma (*shape-preserving*), ou seja, a partir de duas imagens de um objecto

esta técnica produz uma imagem intermédia que representa uma nova vista do mesmo objecto. Assim, é possível criar a ilusão de que os objectos se movem rigidamente (descrevendo movimentos de rotação e translação a 3 dimensões) entre as posições das duas imagens, como se ilustra na Figura 3.5:

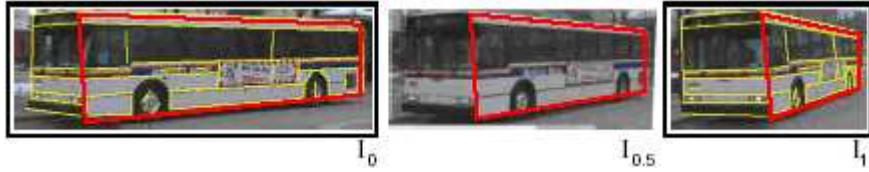


**Figura 3.5 - Síntese da imagem intermédia  $I_{0.5}$ , entre duas vistas distintas do autocarro ( $I_0$  e  $I_1$ ), através da técnica *View-Morphing*, [Seitz\_97].**

As imagens apresentadas pelos autores desta técnica, Seitz e Dyer, levam-nos a concluir que os resultados obtidos através da mesma são mais convincentes e realistas que os resultados alcançados através do método de *Image-Morphing*. Por este motivo, neste trabalho, adoptou-se a técnica *View-Morphing* para sintetizar vistas frontais.

### 3.2.1 Vistas Paralelas

De um ponto de vista geométrico, o processo de formação de uma imagem captada por uma câmara, isto é, o processo de transformação de pontos de um sistema 3D em pontos de um sistema 2D, é representado por operações de projecção (transformações projectivas ou mapas de projecção). Uma transformação projectiva consiste numa aplicação linear que transforma pontos de um espaço vectorial  $P^m$  em pontos do espaço vectorial  $P^n$  (com  $m \leq n$  e  $m, n \in \mathbb{N}$ ). O modelo geométrico de uma câmara representa-se por uma matriz de dimensão  $3 \times 4$ , com a seguinte forma:

$$\Pi = [\mathbf{H} | -\mathbf{H}\mathbf{C}] \quad (\text{Eq. 3.31})$$

onde  $\mathbf{C}$  representa a posição euclidiana do centro óptico da câmara e  $\mathbf{H}$  é uma matriz quadrada de dimensão 3 que especifica a posição e a orientação da imagem da câmara (com respeito ao sistema de coordenadas do mundo), [Seitz\_97].

Suponha-se que se dispõe de duas vistas obtidas com duas câmaras, com eixos ópticos paralelos e com distâncias focais  $f_0$  e  $f_1$  (respectivamente), estando posicionadas de tal forma que o centro óptico da primeira,  $\mathbf{C}_0$ , coincide com a origem do sistema de coordenadas do mundo e o centro óptico da segunda encontra-se na posição  $(C_x, C_y, 0)$ .

As matrizes de projecção  $\mathbf{\Pi}_0$  e  $\mathbf{\Pi}_1$  destas câmaras assumem a seguinte representação:

$$\mathbf{\Pi}_0 = \begin{bmatrix} f_0 & 0 & 0 & 0 \\ 0 & f_0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (\text{Eq. 3.32})$$

e

$$\mathbf{\Pi}_1 = \begin{bmatrix} f_1 & 0 & 0 & -f_1 C_x \\ 0 & f_1 & 0 & -f_1 C_y \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (\text{Eq. 3.33})$$

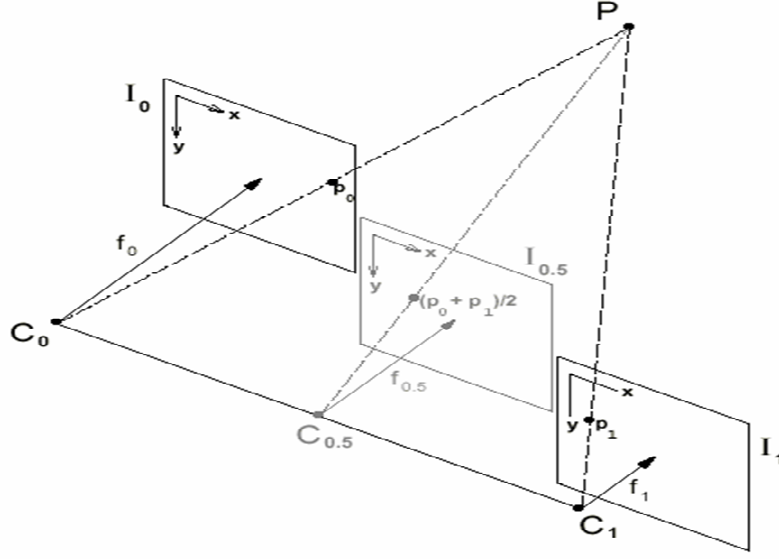


**Figura 3.6 – View-Morphing de vistas paralelas.** A interpolação linear de *pixels* correspondentes em vistas paralelas com planos de imagem  $I_0$  e  $I_1$  produz a imagem  $I_{0.5}$  que representa outra vista paralela do mesmo objecto ou cena do mundo 3D, [Seitz\_97].

Na Figura 3.6, considere-se que o ponto  $P$  da cena, com coordenadas euclidianas  $(X, Y, Z)$ , se expressa, em coordenadas homogêneas, pelo vector coluna  $P = [X \ Y \ Z \ 1]^T$ . Sendo  $p_0 \in I_0$  e  $p_1 \in I_1$  as projecções desse ponto nos planos de imagem  $I_0$  e  $I_1$  (respectivamente), então a interpolação linear entre  $p_0$  e  $p_1$  é dada pela seguinte expressão, [Seitz\_97]:

$$\begin{aligned} (1-s)p_0 + sp_1 &= (1-s)\frac{1}{Z}\Pi_0 P + s\frac{1}{Z}\Pi_1 P \\ \Leftrightarrow (1-s)p_0 + sp_1 &= \frac{1}{Z}\Pi_s P \end{aligned} \quad (\text{Eq. 3.34})$$

com

$$\Pi_s = (1-s)\Pi_0 + s\Pi_1 \quad (\text{Eq. 3.35})$$

e onde o parâmetro  $s \in [0,1]$  define qual a distância da imagem sintetizada às imagens originais.

A técnica *View-Morphing*, dado ser uma transformação que preserva a forma, produz uma nova vista que tem associada a matriz de projecção  $\Pi_s$  que é uma

interpolação linear de  $\Pi_0$  e  $\Pi_1$ . A esta matriz de projecção corresponde o centro da câmara  $C_s$  e a distância focal  $f_s$  dados, respectivamente, por, [Seitz\_97]:

$$C_s = (sC_x, sC_y, 0) \quad (\text{Eq. 3.36})$$

e

$$f_s = (1-s)f_0 + sf_1 \quad (\text{Eq. 3.37})$$

A interpolação de imagens obtidas com câmaras paralelas produz a ilusão simultânea de movimento da câmara ao longo do segmento que une os dois centros ópticos, ou seja, ao longo de  $\overline{C_0C_1}$ , e de *zooming* contínuo. Como referido anteriormente, a interpolação efectuada entre vistas paralelas preserva a forma dos objectos representados. Esta propriedade deve-se à igualdade das terceiras linhas de  $\Pi_0$  e de  $\Pi_1$  [Seitz\_97].

### 3.2.2 Vistas Não Paralelas

Nesta secção descreve-se como se aplica esta técnica na síntese de imagens intermédias de duas vistas não paralelas do mesmo objecto ou cena. Por recurso a uma retro projecção, a síntese de vistas, que preservem a forma dos objectos, a partir de perspectivas não paralelas, reduz-se ao caso de síntese a partir de vistas paralelas abordado anteriormente.

Suponha-se que se dispõe de duas vistas obtidas com duas câmaras, com eixos ópticos não paralelos. Neste caso, a técnica *View-Morphing* começa por efectuar uma retro projecção das vistas  $I_0$  e  $I_1$  num mesmo plano, originando as imagens  $\hat{I}_0$  e  $\hat{I}_1$  (Figura 3.7). Considere-se que  $I_0$  e  $I_1$  possuem as seguinte matrizes de projecção:

$$\Pi_0 = [H_0 | -H_0C_0] \quad (\text{Eq. 3.38})$$

e

$$\mathbf{\Pi}_1 = [\mathbf{H}_1 | -\mathbf{H}_1 \mathbf{C}_1] \quad (\text{Eq. 3.39})$$

onde  $\mathbf{C}_0$  e  $\mathbf{C}_1$  representam os centros ópticos das duas câmaras.

Para simplificar os procedimentos seguintes, o sistema de coordenadas do mundo deve ser escolhido de forma a que os centros ópticos das câmaras recaiam sobre o eixo das abcissas, ou seja, de modo que  $\mathbf{C}_0 = [X_0 \ 0 \ 0]^T$  e  $\mathbf{C}_1 = [X_1 \ 0 \ 0]^T$ .

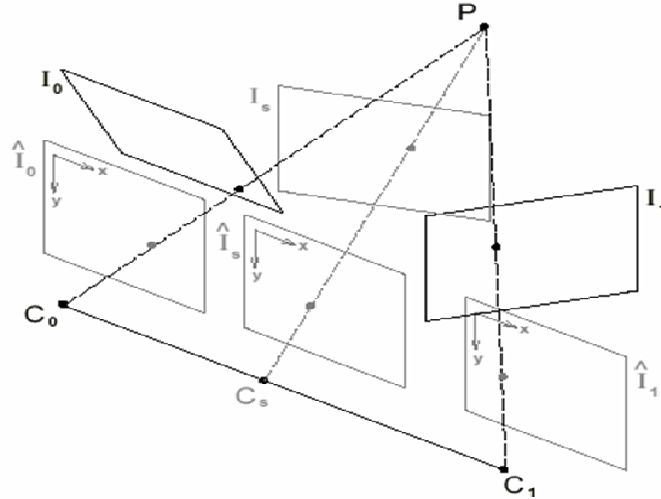


**Figura 3.7 - View-Morphing de vistas não paralelas.** As imagens originais  $I_0$  e  $I_1$  são *prewarped* de modo a formarem as imagens  $\hat{I}_0$  e  $\hat{I}_1$ . A imagem  $\hat{I}_s$  é obtida por interpolação entre as imagens  $\hat{I}_0$  e  $\hat{I}_1$ . Para finalizar a imagem  $\hat{I}_s$  é *postwarped*, originando  $I_s$ , [Seitz\_97].

As vistas de perspectiva na linha  $\overline{\mathbf{C}_0 \mathbf{C}_1}$  podem ser sintetizadas por combinação de operações de retro projecção e de interpolação. Aplicando as transformações projectivas  $\mathbf{H}_0$  a  $I_0$  e  $\mathbf{H}_1$  a  $I_1$ , obtêm-se as imagens paralelas  $\hat{I}_0$  e  $\hat{I}_1$ .

A síntese de uma nova vista  $\hat{I}_s$  entre as imagens  $\hat{I}_0$  e  $\hat{I}_1$  é construída através de interpolação linear de posições, usando a (Eq. 3.34), e intensidades de pontos correspondentes em  $\hat{I}_0$  e  $\hat{I}_1$ .

Para finalizar o processo, resta aplicar a transformação projectiva  $\mathbf{H}_s$  a  $\hat{I}_s$ , obtendo a imagem pretendida  $I_s$ , onde  $\mathbf{H}_s$  é dada por:

$$\mathbf{H}_s = (1-s)\mathbf{H}_0 + s\mathbf{H}_1 \quad (\text{Eq. 3.40})$$

## 3.3 Implementação da técnica *View-Morphing*

Nesta secção descreve-se a implementação da técnica apresentada por Seitz e Dyer, que consiste num algoritmo que os mesmos denominam por algoritmo de 3 passos:

- *Prewarp* – Consiste em transformar as imagens  $I_0$  e  $I_1$  nas imagens paralelas  $\hat{I}_0$  e  $\hat{I}_1$ .
- *Morph* – Este passo tem por objectivo sintetizar, através de interpolações entre as imagens  $\hat{I}_0$  e  $\hat{I}_1$ , a imagem intermédia  $\hat{I}_s$ .
- *Postwarp* – Neste passo final, a imagem  $\hat{I}_s$  é *postwarped*, originando  $I_s$ .

Estes passos encontram-se ilustrados na Figura 3.7.

### 3.3.1 *Prewarp*

A implementação deste passo depende do tipo de projecção que se considera no processo de formação das imagens: perspectiva ou ortográfica.

Atendendo a que, nas imagens adquiridas durante a sequência vídeo, os objectos (neste caso, as faces) têm uma profundidade reduzida comparativamente com a distância a que se encontram da câmara, pode considerar-se que a projecção é ortográfica. Por esta razão, nesta secção apresenta-se apenas a implementação do passo de *Prewarp* no caso de imagens ortográficas. A descrição teórica e o caso geral podem ser consultados em [Seitz\_97] e [Seitz\_96].

Este passo tem por finalidade encontrar as transformações projectivas  $\mathbf{H}_0$  e  $\mathbf{H}_1$ , que, aplicadas, respectivamente, às imagens  $I_0$  e  $I_1$ , as transformam nas imagens  $\hat{I}_0$  e  $\hat{I}_1$  paralelas entre si.

Conhecidas as homografias  $\mathbf{H}_0$  e  $\mathbf{H}_1$  de  $I_0$  e  $I_1$ , respectivamente, o passo de *prewarp* conclui-se aplicando estas matrizes às respectivas imagens. Como se

desconhecem os parâmetros da câmara, também estas matrizes são desconhecidas sendo necessário determiná-las à custa da matriz fundamental.

Considere-se o ponto  $\mathbf{P}$  da cena tridimensional que se representa nas duas imagens  $I_0$  e  $I_1$  por  $\mathbf{p}_0$  e  $\mathbf{p}_1$  respectivamente. Dado um ponto de uma imagem (por exemplo,  $\mathbf{p}_0$ ) a restrição epipolar permite determinar uma recta na outra imagem ( $I_1$ , neste caso) sobre a qual se encontra o ponto correspondente (neste caso,  $\mathbf{p}_1$ ), ([Trucco\_98] – capítulo 7). A seguinte equação traduz a restrição epipolar:

$$\mathbf{p}_1^T \mathbf{F} \mathbf{p}_0 = 0 \quad (\text{Eq. 3.41})$$

onde  $\mathbf{p}_0$  e  $\mathbf{p}_1$  são pontos correspondentes (em coordenadas homogéneas) e  $\mathbf{F}$  é a matriz fundamental. A matriz fundamental é uma matriz 3x3, com característica 2, e encerra toda a informação geométrica que relaciona as duas imagens.

Atendendo a que se considera que a projecção é ortográfica a matriz fundamental assume a seguinte forma, [Seitz\_97] e [Seitz\_96].:

$$\mathbf{F} = \begin{bmatrix} 0 & 0 & a \\ 0 & 0 & b \\ c & d & e \end{bmatrix} \quad (\text{Eq. 3.42})$$

A estimação desta matriz baseou-se no algoritmo sugerido por Hartley e Zisserman [Hartley\_99], que, recorrendo ao método dos mínimos quadrados, necessita apenas de 4 pontos de correspondência entre duas imagens. Como a estimativa desta matriz é bastante instável (o que pode conduzir a erros na síntese das imagens), por forma a aumentar a exactidão no cálculo da mesma foram estabelecidos manualmente 8 pontos de correspondência entre as imagens originais.

As homografias  $\mathbf{H}_0$  e  $\mathbf{H}_1$  são obtidas justamente a partir da relação entre estas e a matriz fundamental, pois quaisquer duas imagens com matriz fundamental  $\mathbf{F}$  podem ser tornadas paralelas escolhendo duas transformações projectivas  $\mathbf{H}_0$  e  $\mathbf{H}_1$  tais que:

$$\left(\mathbf{H}_1^{-1}\right)^T \mathbf{F} \left(\mathbf{H}_0^{-1}\right) = \hat{\mathbf{F}} \quad (\text{Eq. 3.43})$$

onde  $\hat{\mathbf{F}}$  representa a matriz fundamental entre duas imagens paralelas.

Considere-se uma projecção não necessariamente ortográfica. Seja  $E$  um plano paralelo a  $\overline{C_0 C_1}$  (onde  $C_0$  e  $C_1$  representam os centro ópticos das câmaras) e cuja intersecção com as imagens  $I_0$  e  $I_1$  resulta em  $d_0$  e  $d_1$ , respectivamente, como representado na Figura 3.8.

As imagens  $\hat{I}_0$  e  $\hat{I}_1$  podem ser obtidas por projecção de  $I_0$  e  $I_1$  no plano  $E$ , através de uma rotação de cada uma em torno das linhas de intersecção  $d_0$  e  $d_1$ , respectivamente:

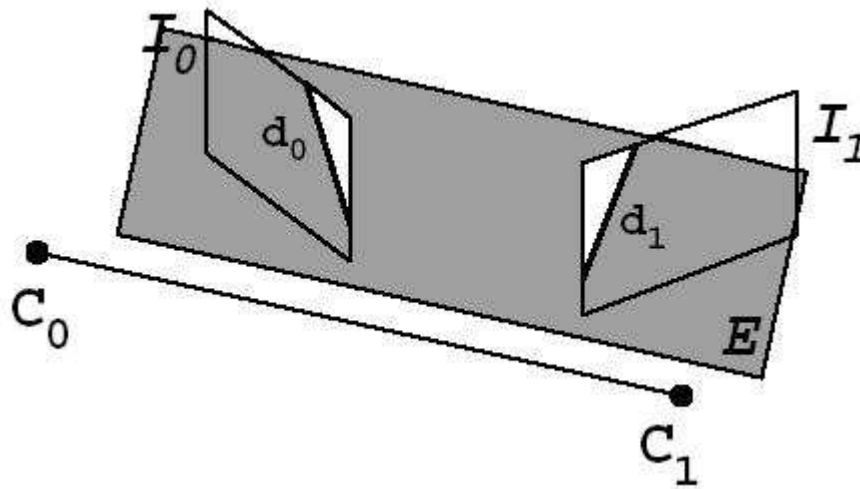


Figura 3.8 – Plano  $E$  que intersecta  $I_0$  e  $I_1$ . [Seitz\_97].

A recta que une  $C_0$  e  $C_1$  intersecta os planos das duas imagens, nos pontos  $e_0$  e  $e_1$ , respectivamente. Estes pontos são designados por epipolos. O epipolo  $e_0$  corresponde à projecção do centro óptico  $C_1$  no plano da imagem  $I_0$  e vice-versa. Como as imagens  $I_0$  e  $I_1$  são ortográficas, os epipolos  $e_0$  e  $e_1$  (que representam a projecção de  $C_1$  em  $I_0$  e vice-versa) são dados por, [Seitz\_97] e [Seitz\_96]:

$$e_0 = \frac{1}{\sqrt{d^2 + c^2}} [-d \ c \ 0]^T \quad (\text{Eq. 3.44})$$

e

$$e_1 = \frac{1}{\sqrt{b^2 + a^2}} [-b \ a \ 0]^T \quad (\text{Eq. 3.45})$$

onde  $a$ ,  $b$ ,  $c$  e  $d$  são os coeficientes da matriz fundamental dada por (Eq. 3.42).

No caso de imagens ortográficas, as projecções  $\mathbf{H}_0$  e  $\mathbf{H}_1$  podem ser encontradas apenas através de uma transformação bidimensional, por forma a alinhar as linhas de varrimento ou *scanlines* das imagens  $\hat{I}_0$  e  $\hat{I}_1$ , tornando-as horizontais. Esta transformação consiste na seguinte rotação:

$$\mathbf{R}_{\phi_i} = \begin{bmatrix} \cos \phi_i & -\sin \phi_i & 0 \\ \sin \phi_i & \cos \phi_i & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (\text{Eq. 3.46})$$

onde  $\phi_i = -\tan^{-1}\left(\frac{e_i^y}{e_i^x}\right)$  e  $e_i^x$  e  $e_i^y$  são as abcissas e as ordenadas dos epipolos, para  $i=1,2$ . Pode acontecer que apesar de alinhadas horizontalmente, as linhas epipolares tenham sentidos opostos; neste caso é necessário proceder a uma rotação adicional de  $180^\circ$ . Esta rotação adicional deverá ser efectuada caso se verifique que  $ab < 0$ , onde  $a$  e  $b$  são coeficientes da matriz fundamental dada pela (Eq. 3.42).

Após esta transformação a matriz fundamental assume a seguinte forma:

$$\tilde{\mathbf{F}} = \mathbf{R}_{\phi_1} \mathbf{F} \mathbf{R}_{-\phi_0} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & a \\ 0 & 1 & b \end{bmatrix} \quad (\text{Eq. 3.47})$$

Para finalizar este processo, a segunda imagem,  $I_1$ , é verticalmente escalada e transladada através da aplicação da seguinte transformação, [Seitz\_97] e [Seitz\_96]:

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -a & -b \\ 0 & 0 & 1 \end{bmatrix} \quad (\text{Eq. 3.48})$$

Concluídas estas transformações, a matriz fundamental entre as imagens *prewarped*  $\hat{I}_0$  e  $\hat{I}_1$  é dada por:

$$\hat{\mathbf{F}} = \mathbf{T}^{-1^T} \mathbf{R}_{\phi_1} \mathbf{F} \mathbf{R}_{-\phi_0} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad (\text{Eq. 3.49})$$

onde  $\hat{\mathbf{F}}$  representa a matriz fundamental entre duas vistas paralelas. Deste modo, e considerando a relação dada por (Eq. 3.43), as homografias a aplicar a  $I_0$  e  $I_1$  de forma a obter as imagens paralelas  $\hat{I}_0$  e  $\hat{I}_1$ , expressam-se como:

$$\mathbf{H}_0 = \mathbf{R}_{\phi_0} \quad (\text{Eq. 3.50})$$

e

$$\mathbf{H}_1 = \mathbf{TR}_{\phi_1} \quad (\text{Eq. 3.51})$$

### 3.3.2 *Morph*

Neste passo sintetiza-se a vista  $\hat{I}_s$  entre as imagens  $\hat{I}_0$  e  $\hat{I}_1$ , a qual fica determinada pelos mapeamentos  $M_0$  e  $M_1$  que especificam completamente a correspondência entre as duas imagens ( $M_0 : \hat{I}_0 \longrightarrow \hat{I}_1$  e  $M_1 : \hat{I}_1 \longrightarrow \hat{I}_0$ ).

Esta síntese consiste essencialmente na deformação das imagens originais e no cálculo da média dos valores relativos à intensidade dos *pixels* na imagem criada. A função de deformação para cada imagem é determinada através de uma interpolação linear entre pontos correspondentes do mapeamento:

$$W_0(\mathbf{p}_0, s) = (1-s)\mathbf{p}_0 + sM_0(\mathbf{p}_0) \quad (\text{Eq. 3.52})$$

e

$$W_1(\mathbf{p}_1, s) = (1-s)M_1(\mathbf{p}_1) + s\mathbf{p}_1 \quad (\text{Eq. 3.53})$$

onde as funções  $W_0$  e  $W_1$  descrevem o deslocamento de cada ponto  $\mathbf{p}_0 \in \hat{I}_0$  e  $\mathbf{p}_1 \in \hat{I}_1$  em função do parâmetro  $s \in [0,1]$  que permite definir qual a distância da imagem sintetizada  $\hat{I}_s$  às imagens  $\hat{I}_0$  e  $\hat{I}_1$ . A imagem sintetizada  $\hat{I}_s$  caracteriza-se por possuir uma matriz de projecção dada por  $\hat{\mathbf{\Pi}}_s = (1-s)\hat{\mathbf{\Pi}}_0 + s\hat{\mathbf{\Pi}}_1$ , sendo  $\hat{\mathbf{\Pi}}_0$  e  $\hat{\mathbf{\Pi}}_1$  as matrizes de

projecção de  $\hat{I}_0$  e  $\hat{I}_1$ , respectivamente, e relaciona-se com as imagens que lhe deram origem através da seguinte expressão:

$$\hat{I}_s(i, j) = (1-s)\hat{I}_0(i, j) + s\hat{I}_1(i, j) \quad (\text{Eq. 3.54})$$

onde  $\hat{I}_s(i, j)$ ,  $\hat{I}_0(i, j)$  e  $\hat{I}_1(i, j)$  representam a intensidade do *pixel*  $(i, j)$ .

A rotina de *morph* utilizada foi desenvolvida por Zhang e Dong [Zhang\_02] e baseia-se no algoritmo de Beier-Neely [Beier\_92], denominado *Field Morphing*. Neste *morph* são seleccionados manualmente pontos característicos correspondentes nas duas imagens. O valor dos *pixels* resultante do *morph* é controlado por estes pontos característicos, sendo o valor de cada *pixel* da imagem resultado calculado consoante as posições relativas destes com respeito às características seleccionadas.

### 3.3.3 Postwarp

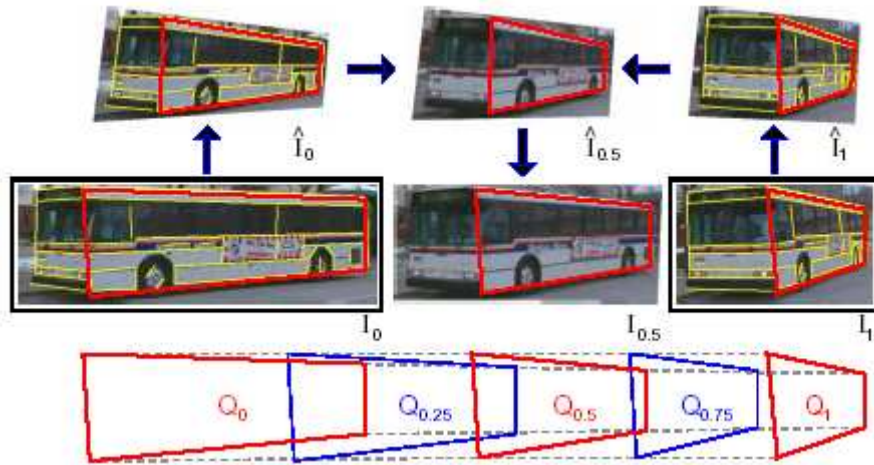


Figura 3.9 - Exemplo do procedimento da técnica *View-Morphing*: selecção de um conjunto de características (linhas amarelas) nas imagens  $I_0$  e  $I_1$ ; utilização destas características para obter as imagens *prewarped*  $\hat{I}_0$  e  $\hat{I}_1$ . A imagem intermédia  $\hat{I}_{0.5}$  é obtida por interpolação entre as imagens  $\hat{I}_0$  e  $\hat{I}_1$ . Para finalizar a imagem  $\hat{I}_s$  é *postwarped*, originando  $I_s$ , [Seitz\_97].

Para finalizar o processo resta aplicar a transformação projectiva  $\mathbf{H}_s$ , dada pela (Eq. 3.40), a  $\hat{I}_s$ , obtendo a imagem pretendida  $I_s$ , sendo  $\mathbf{H}_s$  dada por:

$$\mathbf{H}_s = (1-s)\mathbf{H}_0 + s\mathbf{H}_1 \quad (\text{Eq. 3.55})$$

A sequência de operações anteriormente descrita, e ilustrada na Figura 3.9, permite produzir uma imagem  $I_s$  que corresponde a uma vista com matriz de projecção  $\mathbf{\Pi}_s$ .

# **4 AUTENTICAÇÃO ATRAVÉS DE ICA**

## **4.1 Construção da Base de Dados**

Neste trabalho construiu-se uma base de dados que consiste apenas numa imagem de cada indivíduo, captada em pose frontal e com uma expressão facial neutra, sendo a principal vantagem desta a diminuição de memória ocupada e do tempo de reconhecimento, relativamente à prática comum de armazenar múltiplas vistas da cada indivíduo. A simplicidade desta base de dados recai também no facto de, sendo apenas necessária uma pose de cada indivíduo, o mesmo não precisa de expressar emoções, nem de efectuar trocas de roupa ou de se apresentar com diferentes “adereços”, tais como óculos, barba, chapéu, maquilhagem, entre outros; deste modo o processo de inserção de um determinado indivíduo nesta base de dados torna-se menos incómodo para este e fica concluído em poucos minutos.

Porém, a maioria das bases de dados utilizadas no âmbito da identificação de faces apresenta várias imagens representativas de cada indivíduo (imagens onde o indivíduo se apresenta com diferentes poses, diferentes expressões faciais e com a

presença de “adereços”), o que dificulta a comparação dos resultados obtidos nesta dissertação com trabalhos de outros autores.

Sendo o foco deste trabalho a autenticação de faces, esta é efectuada comparando várias imagens decimadas de uma sequência vídeo de determinado indivíduo com a imagem referente a este na base de dados. Desta forma, o facto da base de dados conter apenas uma imagem de cada indivíduo facilita o processo, no sentido em que este é efectuado de forma mais rápida, pois só há a necessidade de efectuar um *match* de cada imagem da sequência vídeo com uma imagem da base de dados.

Neste trabalho, por forma a validar resultados, antes de efectuar a autenticação de faces precedeu-se ao reconhecimento, confrontando cada imagem da sequência vídeo com todas as imagens da base de dados. Este processo permitiu (após analisadas as identificações e as distâncias de identificação) estabelecer um valor de *threshold* (ou limiar) abaixo do qual as imagens serão autenticadas.

A captação das imagens que integram a base de dados foi efectuada sob condições definidas cuidadosamente, nomeadamente foi exigida a pose frontal dos indivíduos perante a câmara e que os mesmos se encontrem todos à mesma distância desta. As imagens foram adquiridas em tons de cinzento e sendo o fundo de tom cinzento claro e tão uniforme quanto possível. Assim, cada imagem é representada por uma matriz na qual cada elemento revela o valor do *pixel* da imagem na escala de tons de cinzento, ou seja, os elementos da matriz assumem valores inteiros que vão desde o 0 (preto) até ao 255 (branco), na codificação habitual de 8 bits.

Com a finalidade de comparar resultados e avaliar o benefício da síntese de imagens para identificação e autenticação, as imagens extraídas da sequência vídeo foram também confrontadas com as imagens em pose da base de dados.

## 4.2 Síntese de Vistas Frontais das Imagens seleccionadas a partir da Sequência Vídeo

Na aquisição das sequências vídeo a confrontar com as imagens da base de dados foi pedido aos indivíduos que se deslocassem desde um ponto fixo até à câmara de uma forma natural e sem atentarem na mesma.

Da sequência vídeo foram extraídas várias imagens, de dimensão  $640 \times 480$  *pixels*. Por duas razões a qualidade destas imagens é inferior à desejável. Em primeiro lugar, dado que estas imagens representam cenas dinâmicas, verifica-se o efeito *motion blur* causado pelo movimento dos objectos. O segundo motivo, prende-se com o facto de nas imagens mais afastadas da câmara as faces serem representadas com menor resolução. Para contornar estas limitações foram seleccionadas apenas 3 imagens de cada indivíduo (em geral da parte final da sequência), tendo em conta a qualidade das mesmas e diferentes posições do indivíduo perante o eixo óptico (uma pose com rotação, perante o eixo óptico, à esquerda, uma pose com rotação, perante o eixo óptico, à direita e uma pose com o menor ângulo de rotação possível, como se ilustra na Figura 4.1). Depois de escolhidas, foi necessário recolher apenas a informação relativa à face, para que somente esta seja analisada, descartando a informação exterior ao indivíduo a ser reconhecido, para que esta não influencie as tomadas de decisão relativas ao reconhecimento e à autenticação. Após detecção manual das faces, foi efectuado um recorte das imagens e redimensionamento das mesmas para a dimensão  $128 \times 128$  *pixels*, utilizando uma interpolação bicúbica. Assim, após aquisição da sequência vídeo, as imagens foram seleccionadas e foram submetidas ao processo que se apresenta na Figura 4.1.

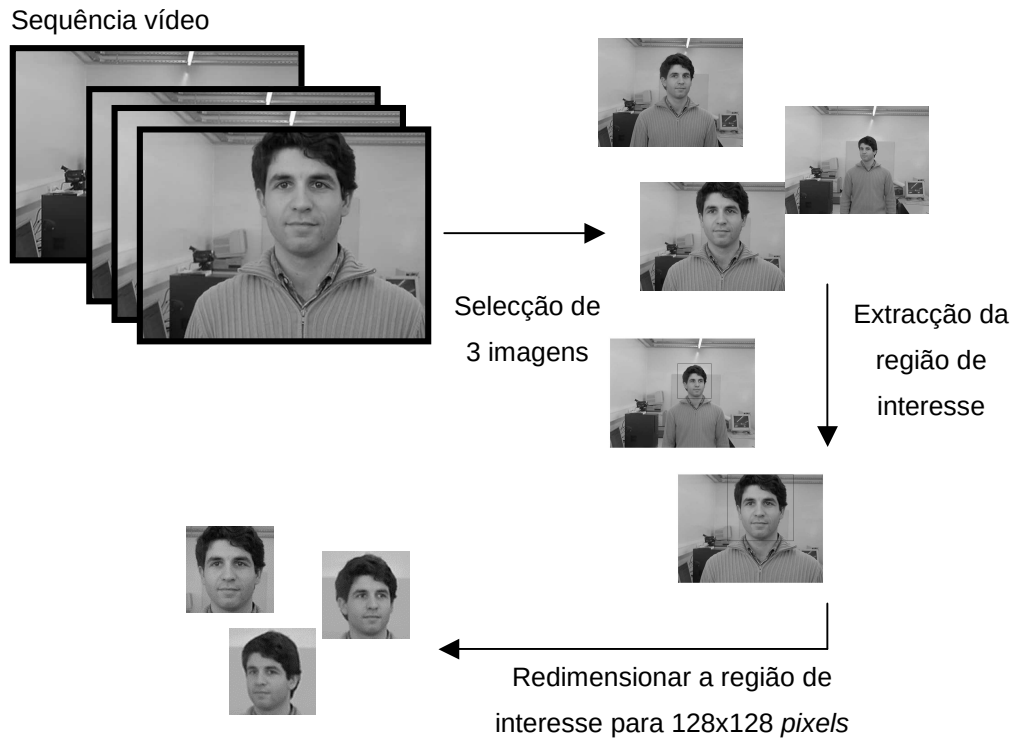


**Figura 4.1 - Processo de selecção, extração da região de interesse e redimensionamento das imagens.**

Para testar a versatilidade do método e a sua aplicabilidade em situações reais, teve-se o cuidado de que as imagens extraídas da sequência vídeo fossem suficientemente distintas entre si, procurando escolher para cada indivíduo imagens em poses com diferentes ângulos em relação ao eixo óptico da câmara. Como as imagens que constituem a base de dados têm uma pose frontal perante o eixo óptico da câmara, recorreu-se à síntese de imagens frontais a partir de vistas laterais. Para cada uma das 3 imagens seleccionadas, determinou-se a sua imagem verticalmente simétrica, como exemplificado na Figura 4.2.

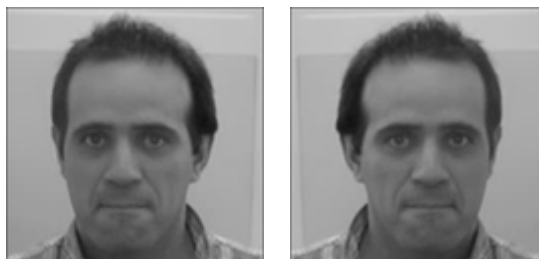


**Figura 4.2 - Imagens simétricas.**

Para sintetizar vistas frontais recorreu-se à técnica *View-Morphing*, referida na secção 3.2. Esta técnica recorre à interpolação geométrica de imagens para, através de

duas imagens de poses diferentes do mesmo objecto e a partir de uma correspondência entre os *pixels* das duas imagens, criar uma nova vista intermédia:

$$I_s = (s - 1)I_0 + sI_1 \quad (\text{Eq. 4.56})$$

onde o parâmetro  $s \in [0,1]$  define qual a distância da imagem sintetizada às imagens originais. Neste trabalho procedeu-se apenas à síntese de uma vista frontal a partir das imagens seleccionadas da sequência vídeo e das suas simétricas, ou seja, o parâmetro  $s$  assume o valor 0.5. Para efectuar a síntese de imagens é necessário estabelecer pontos de correspondência; para que estes não recaiam sobre uma zona uniforme da face foram, em todas as imagens, estabelecidos apenas 8 pontos de correspondência:

- Cantos interiores das sobrancelhas
- Centro de cada olho
- Centro de cada narina
- Cantos dos lábios

A Figura 4.3 ilustra a síntese de vistas intermédias, para diferentes valores do parâmetro  $s$ , através da técnica *View-Morphing* para dois indivíduos:

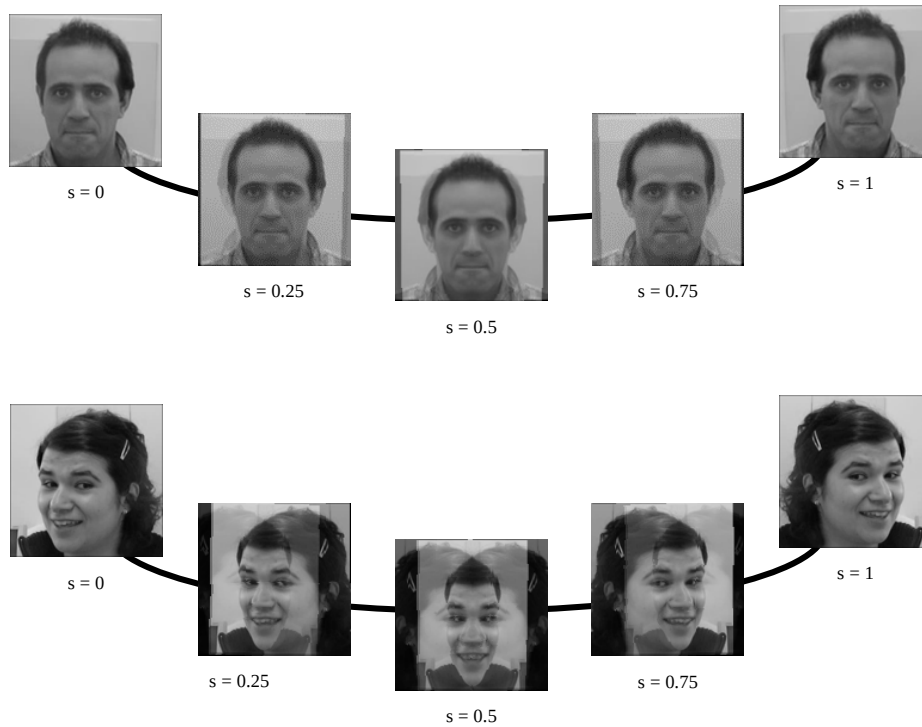


**Figura 4.3 - Síntese de imagens intermédias, através da técnica *View-Morphing*, dependente da escolha do parâmetro  $s$ .**

A segunda sequência da Figura 4.3 ilustra a razão pela qual as vistas escolhidas da sequência vídeo não podem apresentar um ângulo de rotação muito elevado em relação ao eixo óptico da câmara. Como se pode verificar, em qualquer uma das imagens sintetizadas, apenas a zona relativa aos olhos e boca (zona onde foram definidos pontos de correspondência) apresenta uma definição razoável; fora desta zona verifica-se a existência de “fantasmas”, ou seja, vê-se presente a informação de cada uma das imagens utilizadas para a síntese. A existência de “fantasmas” nas imagens sintetizadas poderia ser contornada aumentando o número e distribuição de pontos de correspondência estabelecidos. Como referido na secção 3.1, a tarefa de estabelecer um grande número de pontos de correspondência entre imagens faciais, tendo o cuidado de que os mesmos não recaiam sobre zonas uniformes da face, torna-se difícil e morosa. A morosidade deste processo talvez pudesse ser ultrapassada com a utilização de um algoritmo de detecção automática de pontos correspondentes (esta possibilidade considera-se além dos limites deste trabalho, não tendo por isso sido explorada).

Comparando as imagens das duas sequências da Figura 4.3 é evidente que a qualidade das faces sintetizadas na primeira sequência é melhor que na segunda. Com imagens originais que possuem um menor ângulo de rotação em relação ao eixo óptico

da câmara obtêm-se imagens sintetizadas com menos artefactos devidos ao efeito “fantasma”.

Por este motivo, não serão utilizadas para a síntese de vistas frontais imagens com rotações elevadas, como acontece com as imagens da última sequência da Figura 4.3.

## 4.3 Pré-Processamento das Imagens

Neste trabalho compararam-se resultados de identificação e autenticação para imagens seleccionadas a partir de uma sequência vídeo e para imagens frontais sintetizadas através da técnica *View-Morphing*; deste modo as imagens que formam o conjunto de treino (imagens da base de dados) e as imagens que formam os dois conjuntos de teste considerados foram sujeitas às mesmas transformações de pré-processamento.

As imagens adquiridas para a construção da base de dados tinham uma dimensão de  $2288 \times 1712$  *pixels*. Sujeitar as imagens com esta dimensão ao método ICA além de tornar o processo moroso poderia conduzir a uma perda significativa no desempenho e resultados atingidos. De facto, a elevada dimensão dos dados como também a informação exterior à face do indivíduo poderiam influenciar a identificação e posterior autenticação das faces. Para reduzir a dimensão das imagens e eliminar a informação não específica à face procedeu-se manualmente à identificação da zona da face, efectuando um recorte (*crop*) da mesma e redimensionando-as para  $128 \times 128$  *pixels*, utilizando uma interpolação bicúbica. As imagens seleccionadas a partir da sequência vídeo foram também recortadas e redimensionadas para  $128 \times 128$  *pixels*, conforme referido na secção 4.2.

O objectivo de aumentar a qualidade das imagens e diminuir as discrepâncias entre as mesmas conduziu à necessidade de proceder a um pré-processamento de todas

as imagens, sujeitando as mesmas a transformações, tais como a sua normalização geométrica e fotométrica.

A diminuição do efeito das variações da luminosidade das imagens adquiridas foi efectuada através da equalização do histograma das imagens, produzindo assim uma gradação mais normalizada dos níveis de cinzento.

Para finalizar o pré-processamento das imagens e com a finalidade de corrigir pequenas distorções geométricas existentes efectuou-se uma normalização geométrica. Dado que não se dispõe de informação tridimensional relativa a cada face representada nas imagens da base de dados e da sequência vídeo, a transformação geométrica foi efectuada apenas no plano.

Como se pode verificar na Figura 4.3, a existência de “fantasmas” nas vistas sintetizadas induziu a que as mesmas fossem recortadas, retirando as primeiras 23 colunas de cada imagem e as últimas 24 colunas, obtendo imagens de dimensão  $128 \times 81$  *pixels* (note-se que após a síntese das imagens estas são verticalmente centradas, resultado imediato do facto de se utilizarem imagens simétricas para efectuar a síntese). Para poderem ser comparadas com as imagens sintetizadas, as imagens da base de dados e as imagens seleccionadas a partir da sequência vídeo (originais), após transformação geométrica, foram também recortadas de forma a apresentarem  $128 \times 81$  *pixels*.

### 4.3.1 Transformação Geométrica

O objectivo da normalização geométrica das imagens é corrigir pequenas distorções geométricas, diminuindo a variação entre as representações dos indivíduos. A normalização geométrica a que se submeteram as imagens deste trabalho consistiu numa transformação geométrica que permite obter a representação das faces num sistema plano de coordenadas fixas, definido a partir de três pontos não colineares. Para definir o novo sistema de coordenadas, escolheram-se para pontos de controlo os pontos centrais de cada olho e o ponto médio do segmento que une os cantos dos lábios, com as seguintes posições:

- Ponto central do olho direito: (48, 60)
- Ponto central do olho esquerdo: (80, 60)

- Ponto médio dos lábios: (64, 100)

Esta transformação geométrica facilita a comparação entre imagens, pois as representações obtidas através desta transformação têm características de posicionamento e escala da face iguais em relação à dimensão das imagens.

A transformação geométrica fica definida a partir da identificação da correspondência dos pontos de controlo nas imagens originais e no referencial fixo.

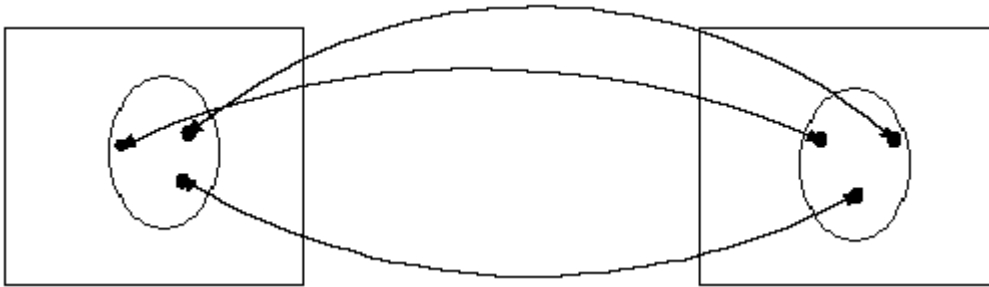


**Figura 4.4 - Correspondência dos pontos de controlo entre o referencial original (esquerda) e o referencial fixo (direita), Adaptado de [Silva\_04].**

Após definida a transformação geométrica, é necessário proceder ao cálculo da intensidade de cada *pixel* da imagem no referencial fixo (com excepção dos pontos de controlo que mantêm o valor original). O cálculo do nível de cinzento de cada *pixel* é efectuado através de métodos de interpolação dos valores de intensidade dos *pixels* originais correspondentes, como interpolação pelo vizinho mais próximo, interpolação bilinear e interpolação bicúbica.

A suavidade do mapeamento dos níveis de cinzento, a simplicidade de implementação computacional, e a rapidez de execução fundamentam que a escolha do método de interpolação tenha recaído sobre a interpolação bilinear.

### 4.3.2 Transformação Fotométrica

O histograma de uma imagem descreve a distribuição estatística dos níveis de cinzento em função do número de *pixels* de cada nível. Pode ser estabelecida uma analogia entre o histograma de uma imagem e a função densidade de probabilidade, que é um modelo matemático da distribuição de tons de cinzento de uma classe de imagens.

Considere-se a imagem  $I$  com  $N$  colunas e  $M$  linhas. O histograma da imagem  $H_I$  define-se como:

$$H_I(k) = \frac{n_k}{NM} \quad (\text{Eq. 4.57})$$

onde  $n_k$  é o número de vezes que o nível de cinzento  $k$  aparece na imagem  $I$ .

A transformação fotométrica ou equalização do histograma das imagens tem por objectivo diminuir o efeito das variações da luminosidade das imagens adquiridas, efectuando um realce de contraste utilizando uma função matemática denominada transformação radiométrica. Esta função consiste em mapear as variações dentro do intervalo original de tons de cinzento, redistribuindo-os para um outro intervalo especificado, produzindo assim uma gradação mais normalizada dos níveis de cinzento, através da expansão do intervalo original de níveis de cinza da imagem original.

A transformação de histograma aplicada às imagens adquiridas consistiu em transformar a intensidade da imagem original numa imagem com 64 níveis de cinzento de forma a uniformizar o histograma desta.

A Figura 4.5 apresenta uma imagem do conjunto de teste e uma do conjunto de treino formado pelas imagens extraídas da sequência vídeo (esquerda) e os resultados obtidos após a equalização do histograma das mesmas (direita):



**Figura 4.5- Exemplos da aplicação da equalização do histograma. Esquerda - Imagens originais. Direita - Imagens resultado.**

A Figura 4.6 esquerda e Figura 4.6 direita apresentam os histogramas das imagens originais anteriores (antes e após a equalização do histograma, respectivamente). Na Figura 4.6 central representa-se o gráfico da função de transformação que permite a equalização do histograma.

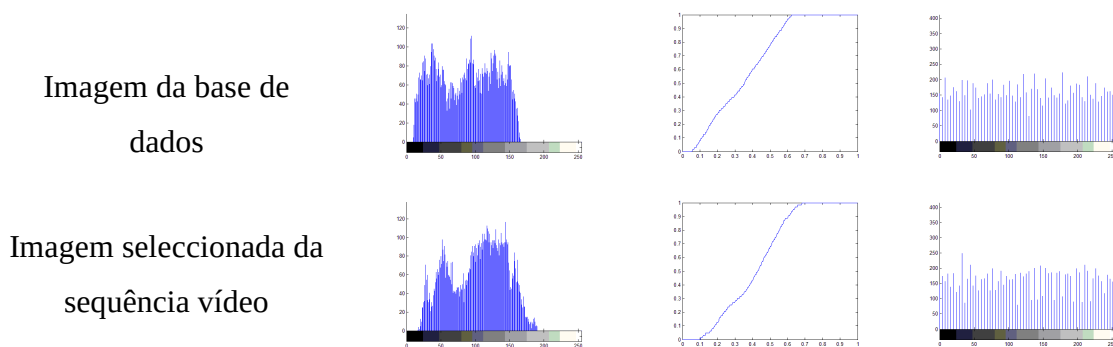


**Figura 4.6 - Exemplos de equalização do histograma. Esquerda - Histograma das imagens originais. Centro - Gráfico das transformações aplicadas. Direita - Histograma das imagens resultado.**

A equalização modifica o histograma da imagem original de tal forma que a imagem transformada tenha um histograma uniforme, ou seja, todos os níveis de cinza devem aparecer na imagem com a mesma frequência. Por observação da Figura 4.6, verifica-se que a obtenção de um histograma perfeitamente uniforme não foi atingida, o que se justifica pela natureza das imagens, nas quais se verifica uma discretização do espaço e da escala de brilho. Contudo, consegue-se melhorar o contraste da imagem,

preenchendo de modo aproximadamente uniforme os 64 possíveis níveis de intensidade distribuídos numa escala de 0 a 255.

## 4.4 Processamento dos Dados

O processamento dos dados consiste em determinar as componentes independentes, encontrando os coeficientes da combinação linear das imagens base estimadas (componentes independentes) e posteriormente em avaliar o reconhecimento e autenticação das imagens.

Ao subespaço próprio gerado pelas componentes independentes atribui-se o nome de espaço ICA e o vector que contém os coeficientes da combinação linear designa-se por representação ICA da imagem. Por aplicação da técnica ICA ao conjunto de dados que constituem o conjunto de treino, obtêm-se as componentes independentes e respectivas representações ICA para cada imagem.

Para avaliar o reconhecimento e autenticação das imagens recorre-se, através de uma medida de distância, à comparação das respectivas representações ICA.

### 4.4.1 Pré-Processamento

Segundo Hyvärinen *et al.*, [Hyvärinen\_01], para garantir maior convergência do algoritmo utilizado, é necessário proceder ao pré-processamento dos dados dos quais se pretende obter a representação ICA.

A centralização dos dados e o branqueamento dos mesmos são tipos de pré-processamento que devem ser efectuados.

O processo de centrar os dados corresponde a calcular uma estimativa da média,  $\hat{\mu}$ , dos dados e posteriormente subtrair a estimativa calculada aos mesmos, tornando assim nula a média dos dados da imagem:

$$\mathbf{X}_C = \mathbf{X} - \hat{\mu} \quad (\text{Eq. 4.58})$$

Após centrar os dados  $\mathbf{X}_C$  é o conjunto de dados a utilizar. Para simplificar a notação  $\mathbf{X}_C$  passa a ser denotado por  $\mathbf{X}$ .

A finalidade do branqueamento ou esferização dos dados é a normalização destes relativamente à variância. A matriz de branqueamento é definida à custa das matrizes de vectores próprios e valores próprios da matriz de covariâncias:

$$\mathbf{V} = \mathbf{E} \mathbf{D}^{-\frac{1}{2}} \mathbf{E}^T \quad (\text{Eq. 4.59})$$

onde  $\mathbf{E}$  é a matriz dos autovectores e  $\mathbf{D}$  é a matriz dos autovalores da matriz de covariâncias  $\Sigma$ .

A matriz de dados  $\tilde{\mathbf{X}}$  com média zero e matriz de covariâncias igual à identidade é obtida pela seguinte transformação:

$$\tilde{\mathbf{X}} = \mathbf{V} \mathbf{X} = \mathbf{V} \mathbf{A} \mathbf{s} = \tilde{\mathbf{A}} \mathbf{s} \quad (\text{Eq. 4.60})$$

Dado que a matriz  $\tilde{\mathbf{A}}$  é ortogonal, este processo reduz a procura da matriz de separação ao espaço das matrizes ortogonais, reduzindo a complexidade do problema e melhorando o desempenho do método ICA.

#### 4.4.2 Processamento - Selecção do Método de Estimação, Algoritmo e Função Objectivo Utilizados na Autenticação

Como foi referido na secção 2.4, a escolha da função objectivo e do algoritmo a utilizar na execução do método ICA depende das propriedades estatísticas da primeira e das propriedades algorítmicas do segundo.

Fundamentada nos resultados apresentados e nas considerações tecidas por Hyvärinen *et al.*, [Hyvärinen\_01], a escolha do método de estimação recai sobre o *Método da Maximização da Não Gaussiannidade*. Devido à sua robustez como medida

de não gaussianidade, adoptou-se a aproximação à *neguentropia* através de funções não polinomiais (Eq. 2.24).

Como função de ordem superior à quadrática utilizou-se a função logaritmo do cosseno hiperbólico  $G(z) = \log \cosh(z)$ , pois esta expressão representa a densidade logarítmica de uma distribuição super gaussiana [Hyvärinen\_99], e como função objectivo ou função de custo (definida como a derivada da função  $G$ ) a tangente hiperbólica:  $g(y) = \tanh(y)$ . Propriedades estatísticas como consistência, variância assintótica e robustez, fundamentam esta escolha.

A eficiência, a memória utilizada, a rapidez de convergência e a estabilidade numérica do algoritmo rápido do ponto fixo (*fastICA*) justificam a sua adopção como algoritmo de optimização a utilizar na estimação das componentes independentes. Dado que a precisão da máquina pode conduzir a erros numéricos, as componentes independentes são estimadas simultaneamente, por forma a reduzir a propagação dos erros no decorrer das sucessivas iterações.

O algoritmo *fastICA* é um método computacional altamente eficiente para atingir a estimação de componentes independentes. Usa um esquema de repetição do ponto-fixa desenvolvido em experiências independentes com a finalidade de ser mais rápido que métodos de gradiente convencionais.

A autenticação e reconhecimento das faces são efectuados por comparação das respectivas projecções no espaço ICA. A metodologia aplicada para determinar as representações ICA consiste nas seguintes etapas:

- Definição de um conjunto de treino,  $\mathbf{X}_{\text{treino}}$ , e de um conjunto de teste,  $\mathbf{X}_{\text{teste}}$ .  
Neste trabalho, o conjunto de treino é composto pelas imagens que constituem a base de dados. Para comparar com este, foram utilizados dois conjuntos de teste, formados pelas imagens retiradas da sequência vídeo e pelas imagens sintetizadas através da técnica *View-Morphing*, originando os conjuntos  $\mathbf{X}_{\text{teste}_1}$  e  $\mathbf{X}_{\text{teste}_2}$ , respectivamente.
- Utilização do conjunto de treino para estimar as componentes independentes por aplicação da técnica ICA, recorrendo ao método de estimação, à função objectivo e ao algoritmo de optimização referidos no início desta secção.

- Definição da matriz de coeficientes ICA,  $\mathbf{B}_{\text{treino}}$ , que, considerando a equação (Eq. 2.2), permite obter  $\mathbf{X}_{\text{treino}}$  em função das componentes independentes estimadas  $\hat{\mathbf{s}}$ .
- Definição dos vectores de coeficientes ICA,  $\mathbf{B}_{\text{teste}}$  ( $\mathbf{B}_{\text{teste}_1}$  e  $\mathbf{B}_{\text{teste}_2}$ ), referentes ao conjunto  $\mathbf{X}_{\text{teste}}$  ( $\mathbf{X}_{\text{teste}_1}$  e  $\mathbf{X}_{\text{teste}_2}$ ), que correspondem à projecção de cada imagem de  $\mathbf{X}_{\text{teste}}$  no espaço ICA gerado pelas componentes independentes estimadas  $\hat{\mathbf{s}}$  e que são obtidos através da relação:

$$\mathbf{X}_{\text{teste}} = \mathbf{B}_{\text{teste}} \hat{\mathbf{s}} \Rightarrow \mathbf{B}_{\text{teste}} = \mathbf{X}_{\text{teste}} \hat{\mathbf{s}}^T (\hat{\mathbf{s}} \hat{\mathbf{s}}^T)^{-1} \quad (\text{Eq. 4.61})$$

A escolha da arquitectura para construção dos conjuntos de treino e de teste prende-se com a dimensão das imagens. O conjunto de treino é formado por 22 imagens, sendo cada uma constituída por  $128 \times 81 = 10368$  *pixels*, o que implica que as matrizes geradas durante a aplicação da técnica ICA tenham dimensões de  $22^2$  ou  $10368^2$ , consoante se considere a arquitectura I ou a arquitectura II. Assim, a fim de reduzir o esforço computacional, a escolha recaiu sobre a arquitectura I.

### 4.4.3 Pós-Processamento: Escolha da medida de distância entre imagens

A avaliação da diferença entre representações ICA de duas imagens, uma pertencente ao conjunto  $\mathbf{X}_{\text{treino}}$  e outra a um conjunto de teste é feita através da definição de uma distância entre elas.

Bartlett *et al.*, [Bartlett\_01], sugerem a distância cosseno:

$$d(i, j) = 1 - \frac{\mathbf{X}_{\text{teste}}(j) \cdot \mathbf{X}_{\text{treino}}^T(i)}{\|\mathbf{X}_{\text{teste}}(j)\| \|\mathbf{X}_{\text{treino}}^T(i)\|} \quad (\text{Eq. 4.62})$$

onde  $\mathbf{X}_{\text{teste}}(j)$  e  $\mathbf{X}_{\text{treino}}^T(i)$  representam os vectores linha  $j$  e  $i$  das matrizes  $\mathbf{X}_{\text{teste}}$  e  $\mathbf{X}_{\text{treino}}^T$ , respectivamente.

Assim, a face da imagem  $j$  do conjunto de teste é identificada como a face  $i$  do conjunto  $\mathbf{X}_{\text{treino}}$  se a distância dada por (Eq. 4.62) for suficientemente próxima de zero.

Para proceder à autenticação, a face da imagem  $j$  do conjunto de teste é autenticada como a face  $j$  do conjunto  $\mathbf{X}_{\text{treino}}$  se a distância das imagens extraídas da sequência vídeo relativas ao indivíduo  $j$ , dada por (Eq. 4.62) não ultrapassar determinado limiar de significância (definido após análise dos resultados obtidos para a identificação dos indivíduos).

## 5 RESULTADOS OBTIDOS

No decurso deste projecto foram seguidas e adoptadas diferentes metodologias, tendo por base resultados obtidos que indiciavam novos procedimentos. Por vezes, alguns destes testes revelaram-se pouco interessantes não atingindo os resultados esperados. Durante esta fase de testes, e na tentativa de obter resultados mais fiáveis e seguros no que o concerne à identificação dos indivíduos, tentou-se aumentar a robustez e a consistência da base de dados, que originalmente era constituída por apenas uma vista frontal de cada indivíduo. Desta forma considerou-se um conjunto de treino artificialmente aumentado, constituído por 10 imagens frontais de cada indivíduo, diferenciadas entre si por pequenas translações verticais e horizontais. Porém os testes efectuados demonstraram que os benefícios decorrentes do aumento do número de imagens por indivíduo não compensavam o incremento do esforço e memória computacional, sendo que o modelo ICA definido a partir deste conjunto de treino demorava, por vezes, cerca de 24 horas a ser obtido. Assim, descartou-se esta hipótese, abordando metodologias assentes numa base de dados com apenas uma imagem facial de cada indivíduo.

Estes testes de experimentação conduziram à metodologia descrita no capítulo 2. No presente capítulo apresentam-se e analisam-se os resultados obtidos, demonstrando a aplicabilidade da técnica de Análise de Componentes Independentes ao reconhecimento de imagens. Para averiguar a vantagem da síntese de imagens, abordada no capítulo 3, no reconhecimento de imagens foram efectuados testes que permitem comparar resultados e avaliar o impacto da síntese no reconhecimento. Deste modo, para explorar este recurso foram definidos diferentes conjuntos de imagens: um conjunto de treino e

dois conjuntos de teste. As imagens do conjunto de treino são utilizadas para definir o modelo ICA, ou seja, é com estas imagens que se obtêm os coeficientes que transformam os dados (imagens) do espaço original para o espaço ICA. Os conjuntos de teste permitem avaliar o reconhecimento utilizando para tal uma medida de classificação.

Neste trabalho pretende-se simular uma situação real de autenticação de forma a permitir ou não o acesso de um indivíduo a uma área restrita. Para efectuar a autenticação de um indivíduo que se identifica como tendo acesso a uma determinada área ou serviço é comparada uma representação no espaço ICA de uma imagem (ou mais) deste com o modelo ICA do indivíduo com o qual este se identifica, presente na base de dados. A comparação das representações é efectuada através da medida de reconhecimento definida na secção 4.4.3 e considera-se que um indivíduo é autenticado se a distância entre as representações for inferior a um determinado limiar.

Para simular esta situação, e como referido na secção 4.1, foi construída uma base de dados com apenas uma imagem em pose frontal de 22 indivíduos. Na Figura 5.1 apresentam-se as imagens pose de indivíduos que constituem a base dados. Aquando da construção da base de dados, os indivíduos responderam a uma formulário onde autorizavam ou não a divulgação e o uso das imagens captadas; por esta razão, não é possível apresentar as imagens de todos os 22 indivíduos, cingindo-nos apenas aos que autorizaram que as suas imagens fossem tornadas públicas.



**Indivíduo “ar001”**



**Indivíduo “ar003”**



**Indivíduo “ar004”**



**Indivíduo “ar006”**



**Indivíduo “ar011”**



**Indivíduo “ar012”**



**Indivíduo “ar013”**



**Indivíduo “ar014”**



**Indivíduo “ar015”**



**Indivíduo “ar017”**



**Indivíduo “ar018”**



**Indivíduo “ar019”**



**Indivíduo “ar020”**



**Indivíduo “ar021”**



**Indivíduo “ar022”**

**Figura 5.1 – Imagens, em pose frontal, de alguns indivíduos que constituem a base de dados.**

De cada sequência vídeo captada extraíram-se três imagens faciais, como descrito na secção 4.2. Foi a partir destas imagens que se construíram os dois conjuntos de teste; o primeiro é constituído pelas imagens extraídas da sequência vídeo e o segundo é formado pelas vistas frontais sintetizadas, utilizando a técnica *View-Morphing*, a partir das imagens presentes no primeiro conjunto de teste. A definição do valor de *threshold* ou valor limiar abaixo do qual as imagens são autenticadas foi efectuada experimentalmente, por recurso à identificação de cada indivíduo, confrontando cada imagem sintetizada a partir da sequência vídeo com todas as imagens da base de dados.

Deste modo neste capítulo apresentam-se e analisam-se os resultados obtidos para a identificação e autenticação dos indivíduos, utilizando para o efeito os dois conjuntos de teste referidos. Os resultados obtidos para ambos os conjuntos de teste são ainda comparados de forma a avaliar o benefício da síntese de imagens. Convém lembrar que as imagens foram previamente processadas conforme descrito na secção 4.3.

Os resultados da identificação são avaliados através da taxa de acerto, que representa o quociente entre o número de imagens correctamente identificadas ( $N_{\text{correctos}}$ ) e o número total de imagens de teste ( $N_{\text{total}}$ ) e é dada por:

$$T_{\text{identificação}} = \frac{N_{\text{correctos}}}{N_{\text{total}}} \times 100\% \quad (\text{Eq. 5.63})$$

Assim, a face da imagem  $i$  do conjunto de teste é identificada correctamente como a face  $i$  do conjunto de treino se a distância dada pela (Eq. 4.62) for a mínima e se for suficientemente próxima de zero.

Para proceder à autenticação definiu-se um limiar de significância para cada indivíduo utilizando a distância de cada imagem facial do indivíduo  $i$  extraída da sequência de vídeo a todas as representações do conjunto de treino. Este valor de *threshold*, para cada indivíduo, é dado pela (Eq. 5.64):

$$T_i = \mu_i - \sigma_i \quad (\text{Eq. 5.64})$$

onde

$$\mu_i = \frac{\sum_{k=1}^{N_{\text{images}}} \sum_{j=1}^{N_{\text{ind}}} d_k(i, j)}{N_{\text{images}} \cdot N_{\text{ind}}} \quad \text{e} \quad \sigma_i = \sqrt{\frac{\sum_{k=1}^{N_{\text{images}}} \sum_{j=1}^{N_{\text{ind}}} [d_k(i, j) - \mu_i]^2}{(N_{\text{images}} \cdot N_{\text{ind}} - 1)}} \quad (\text{Eq. 5.65})$$

com  $N_{\text{images}} = 3$  e  $N_{\text{ind}} = 22$ .

Deste modo, se o indivíduo  $i$  se identificar como tal, este é autenticado correctamente como o indivíduo  $i$  se a distância, dada pela (Eq. 4.62), de, pelo menos, duas das três imagens faciais extraídas da sequência vídeo à representação ICA do indivíduo  $i$  for inferior ao valor de *threshold* calculado. É esta distância entre as representações que revela o grau de semelhança entre as duas imagens. Para avaliar os resultados da autenticação definiu-se uma taxa de acerto, que representa o quociente entre o número de indivíduos correctamente autenticados ( $N_{\text{Ind\_correctos}}$ ) e o número total de indivíduos pertencentes à base de dados ( $N_{\text{Ind\_total}}$ ) e traduz-se pela (Eq. 5.66):

$$T_{\text{autenticação}} = \frac{N_{\text{Ind\_correctos}}}{N_{\text{Ind\_total}}} \times 100\% \quad (\text{Eq. 5.66})$$

## 5.1 Comparação entre os Resultados de Identificação das Imagens da Sequência Vídeio e das Imagens Sintetizadas

Para indagar os benefícios da síntese de imagens frontais, confrontaram-se os resultados de identificação, obtidos para os conjuntos de teste constituídos pelas imagens extraídas da sequência vídeo ( $X_{\text{teste}_1}$ ), e pelas vistas frontais sintetizadas, utilizando a técnica *View-Morphing*, a partir das primeiras ( $X_{\text{teste}_2}$ ). Estes resultados são obtidos por comparação das distâncias das representações dos conjuntos de teste

( $X_{\text{teste}_1}$  e  $X_{\text{teste}_2}$ ) às representações de um conjunto de treino constituído pelas imagens captadas em pose frontal, sendo estas recortadas para a dimensão  $128 \times 81$  *pixels* e submetidas a uma equalização do histograma. Na Figura 5.2 mostram-se as imagens, em pose frontal, de alguns dos indivíduos que compõem o conjunto de treino  $X_{\text{treino}_1}$ :

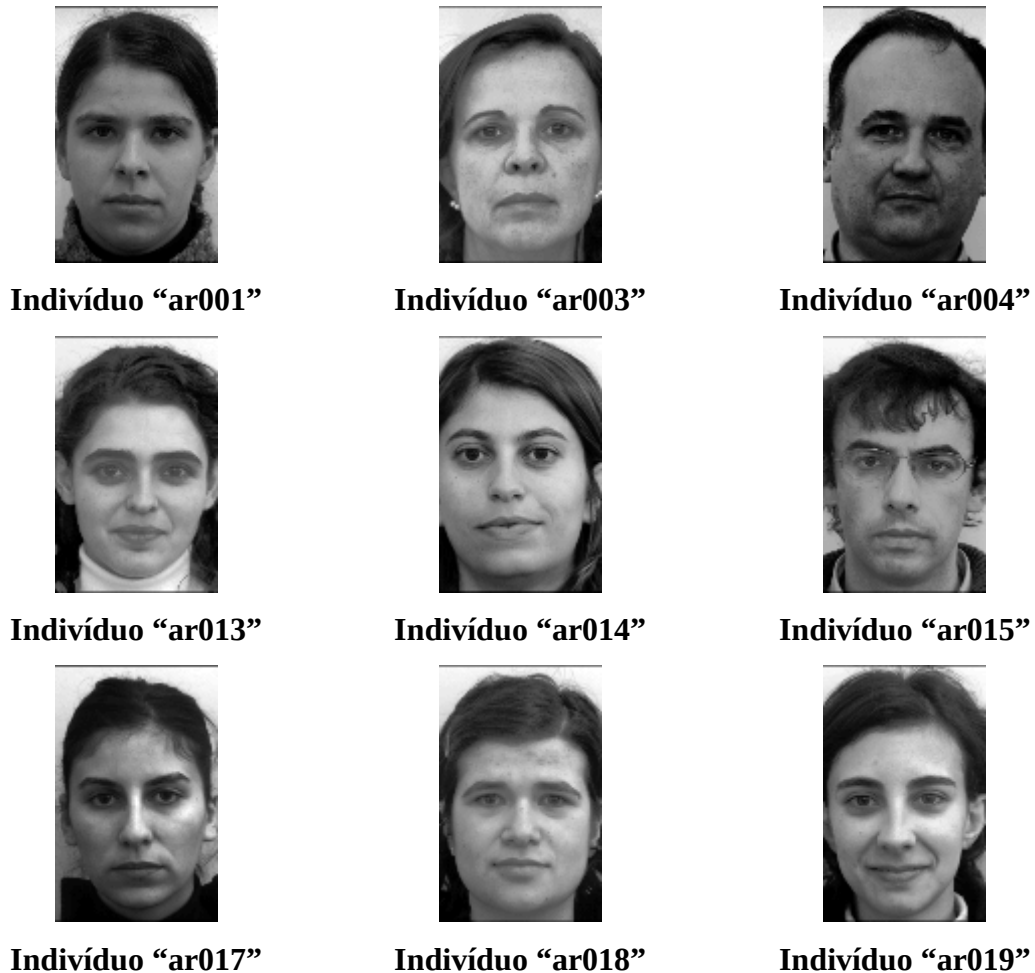


**Figura 5.2 – Imagens, em pose frontal, de alguns indivíduos que constituem o conjunto de treino  $X_{\text{treino}_1}$ .**

As imagens de teste são imagens de dimensão  $128 \times 81$  *pixels* e sem correcção geométrica, sendo apenas efectuada uma equalização das imagens. Nas Figura 5.3 e 5.4 ilustram-se algumas das imagens que constituem os conjuntos de teste  $X_{\text{teste}_1}$  e  $X_{\text{teste}_2}$ , respectivamente.

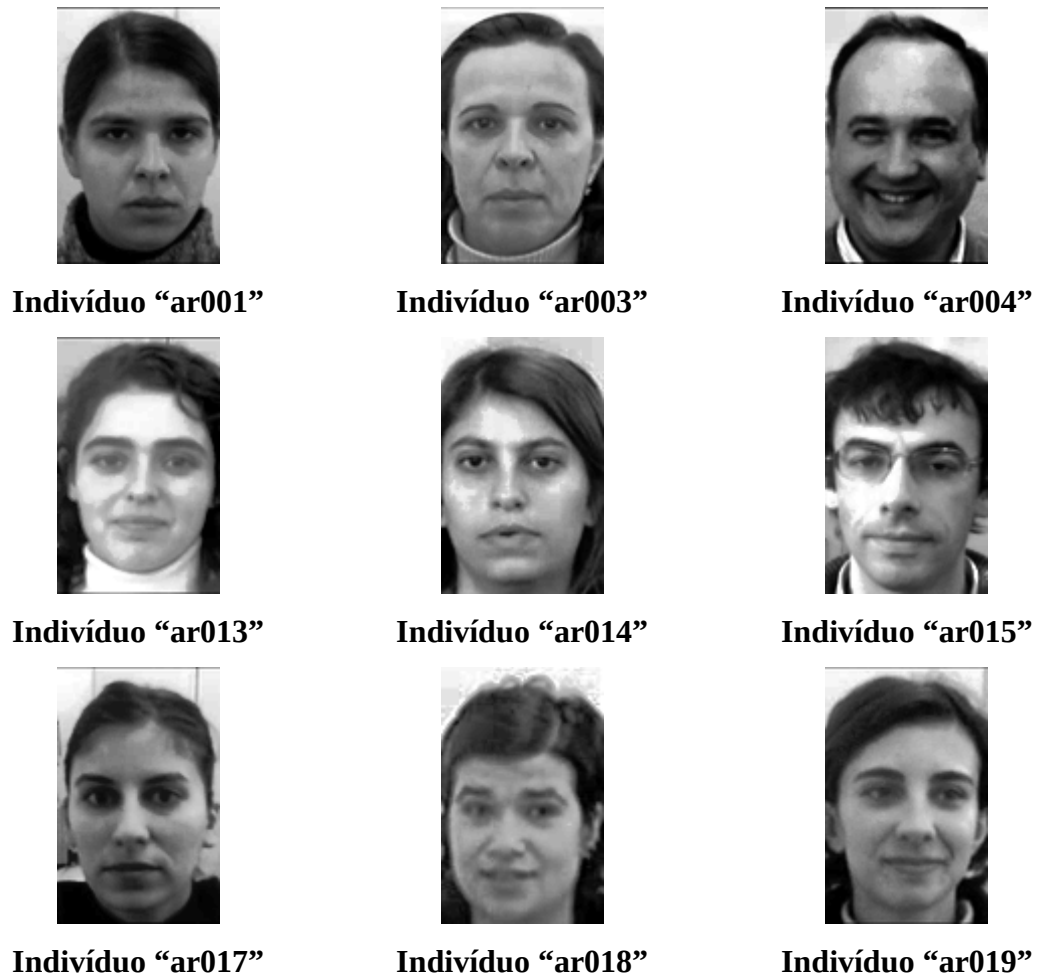


Figura 5.3 – Algumas imagens extraídas da sequência de vídeo e que constituem o conjunto de teste  $X_{\text{teste}_1}$ .

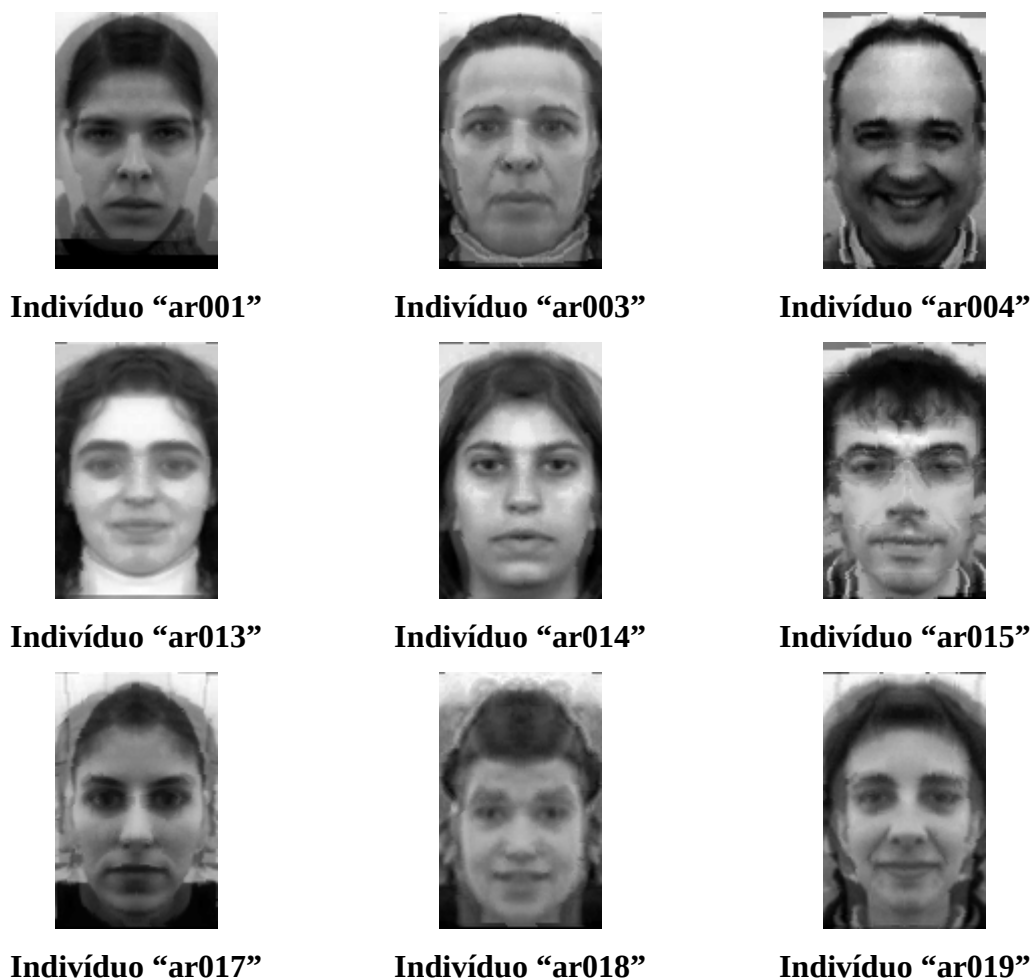


**Figura 5.4 – Algumas imagens sintetizadas, a partir das imagens extraídas da sequência de vídeo, e que constituem o conjunto de teste  $X_{\text{teste}_2}$ .**

A Tabela 5.1 apresenta os resultados obtidos para a identificação dos indivíduos com os dois conjuntos de teste referidos.

**Tabela 5.1 - Taxas de acerto obtidas para a identificação por aplicação da técnica ICA às imagens retiradas da sequência vídeo e às imagens sintetizadas a partir destas.**

|                                                    | Imagens retiradas<br>da sequência vídeo | Imagens<br>sintetizadas      |
|----------------------------------------------------|-----------------------------------------|------------------------------|
| Taxa de identificação - $T_{\text{identificação}}$ | $\frac{32}{66} \times 100\%$            | $\frac{35}{66} \times 100\%$ |

Como se pode verificar, a diferença entre as taxas de acerto é mínima, sendo relativamente mais elevada no caso de se utilizarem as imagens sintetizadas ( $X_{\text{teste}_2}$ ). Porém, no que concerne à comparação entre resultados obtidos com imagens retiradas

da sequência vídeo e imagens sintetizadas a partir destas, deve ter-se em conta que as imagens sintetizadas apresentam uma qualidade inferior às originais. De facto, a presença de artefactos nas imagens sintetizadas, referidos na secção 4.2 como “fantasmas”, contribui não só para a existência de informação errónea como também a existência de mais informação não específica da face. Este problema foi abordado na secção 4.2, onde se referiu que as imagens sintetizadas, devido ao reduzido número de pontos de correspondência e à sua distribuição não uniforme em toda a imagem, apresentam uma qualidade inferior à desejada.

Por este motivo, é injusto, dada a qualidade das imagens sintetizadas, comparar resultados entre as imagens seleccionadas da sequência vídeo e as sintetizadas.

Não obstante a injustiça da comparação efectuada, os resultados obtidos e apresentados nesta secção são os alicerces que nos permitiram decidir pela utilização das imagens sintetizadas em vez das imagens extraídas da sequência vídeo.

## **5.2 Comparação entre os Resultados Sem e Com Correção Geométrica das Imagens**

Para avaliar a necessidade de efectuar ou não a correção geométrica das imagens (referida na secção 4.3), foram construídos um novo conjunto de treino e um novo conjunto de teste. O novo conjunto de treino,  $\mathbf{X}_{\text{treino}_2}$ , é composto pelas mesmas imagens, em pose frontal, que o conjunto de treino  $\mathbf{X}_{\text{treino}_1}$ , porém após correção geométrica. Na Figura 5.5 ilustram-se algumas das imagens presentes no conjunto de treino  $\mathbf{X}_{\text{treino}_2}$ :

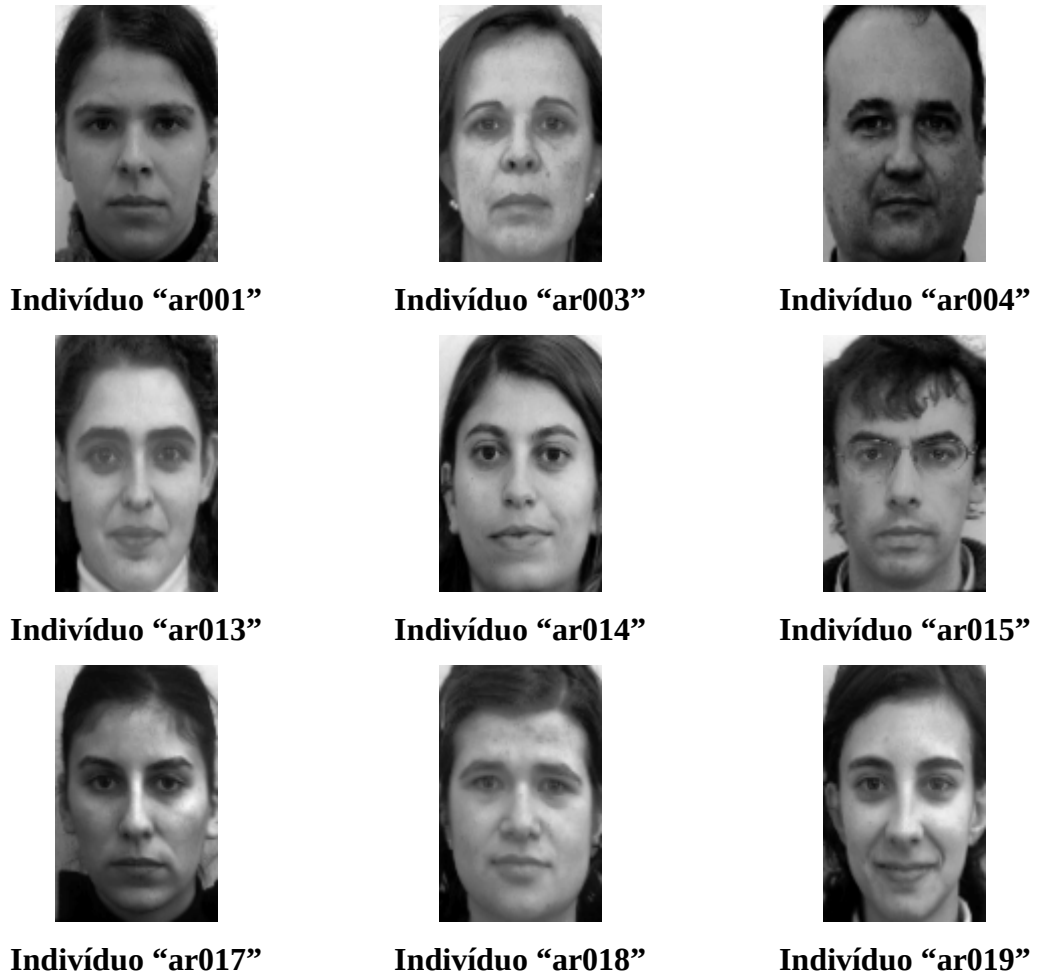


**Figura 5.5 – Imagens, em pose frontal, de alguns indivíduos que constituem o conjunto de treino  $X_{\text{treino}_2}$ .**

Os dois conjuntos de teste considerados são formados pelas imagens sintetizadas através da técnica *View-Morphing*. O primeiro,  $X_{\text{teste}_2}$ , é o conjunto referido na secção 5.1, e é constituído pelas imagens sintetizadas sem correcção geométrica. O segundo conjunto de teste,  $X_{\text{teste}_3}$ , é formado por estas imagens após correcção geométrica. Na Figura 5.6 ilustram-se algumas das imagens presentes no conjunto de teste  $X_{\text{teste}_3}$ :

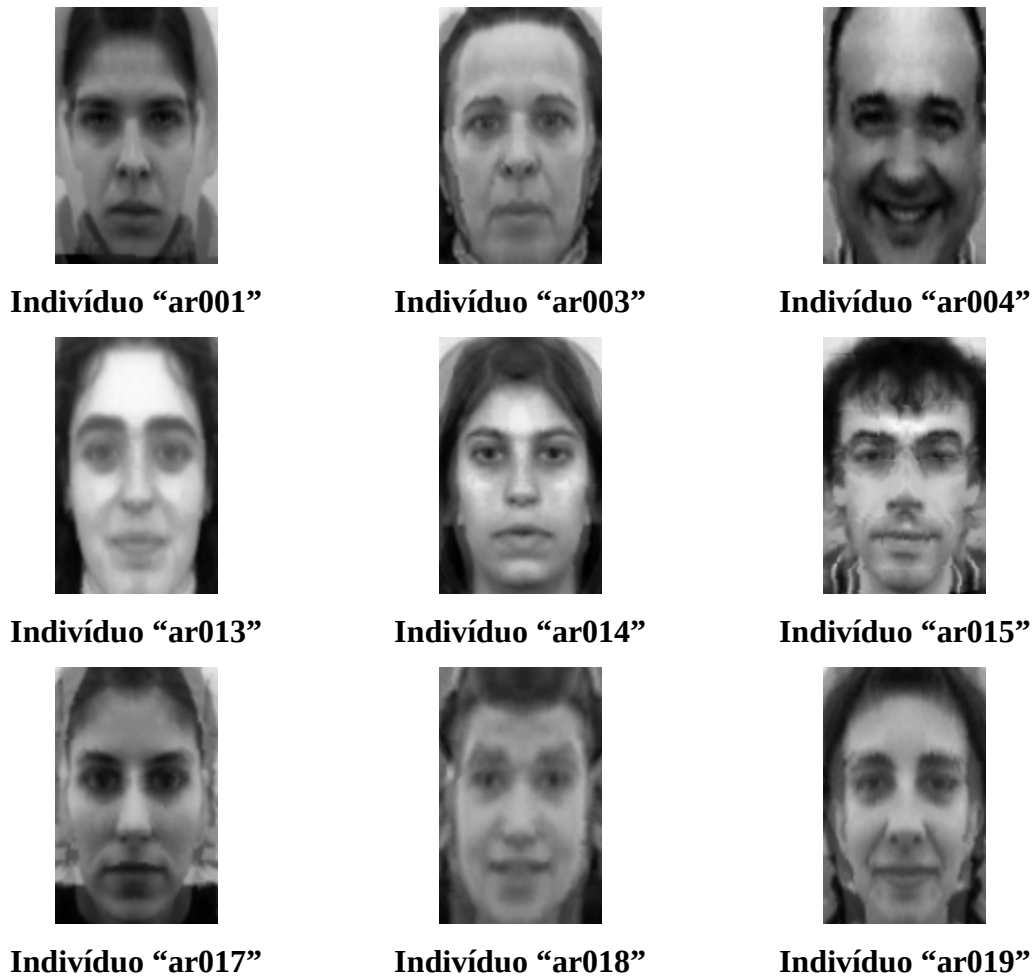


**Figura 5.6 – Algumas imagens sintetizadas e após correcção geométrica e que constituem o conjunto de teste  $X_{teste_3}$ .**

Assim, depois de definidas as representações ICA para os conjuntos de treino  $X_{treino_1}$  e  $X_{treino_2}$ , são comparadas as distâncias entre as representações do conjunto de teste,  $X_{teste_2}$  e as representações do conjunto de treino  $X_{treino_1}$ . Analogamente, comparam-se as distâncias entre as representações do conjunto  $X_{teste_3}$  e as representações de  $X_{treino_2}$ .

Confrontando estes resultados, verifica-se, quando utilizados os conjuntos de treino e de teste formados pelas imagens com correcção geométrica, um acréscimo na taxa de identificação. Na Tabela 5.2 apresentam-se os resultados obtidos, para identificação das imagens, por aplicação da metodologia ICA às imagens sintetizadas sem e após correcção geométrica.

**Tabela 5.2 - Taxas de acerto para a identificação de imagens obtidas por aplicação da técnica ICA às imagens sintetizadas sem e com correcção geométrica.**

|                                                       | Imagens sintetizadas sem<br>correcção geométrica | Imagens sintetizadas com<br>correcção geométrica |
|-------------------------------------------------------|--------------------------------------------------|--------------------------------------------------|
| Taxa de identificação -<br>$T_{\text{identificação}}$ | $\frac{35}{66} \times 100\%$                     | $\frac{40}{66} \times 100\%$                     |

Verifica-se que a correcção geométrica é vantajosa, conduzindo a uma taxa de acerto mais elevada, sendo que deste modo se obtêm mais identificações correctas. Como referido no início deste capítulo, o valor limiar para autenticar um determinado indivíduo é obtido, neste caso, por recurso às distâncias entre as imagens sintetizadas (sem e com processamento) e todas as representações ICA do conjunto de treino ( $X_{\text{treino}_1}$  e  $X_{\text{treino}_2}$ , respectivamente), utilizando a (Eq. 5.64). Na Tabela 5.3 apresentam-se os resultados para autenticação dos dois conjuntos de teste em causa.

**Tabela 5.3 - Taxas de acerto para a autenticação de imagens obtidas por aplicação da técnica ICA às imagens sintetizadas sem e com correcção geométrica.**

|                                                     | Imagens sintetizadas sem<br>correcção geométrica | Imagens sintetizadas com<br>correcção geométrica |
|-----------------------------------------------------|--------------------------------------------------|--------------------------------------------------|
| Taxa de autenticação -<br>$T_{\text{autenticação}}$ | $\frac{15}{22} \times 100\%$                     | $\frac{22}{22} \times 100\%$                     |

Os resultados apresentados permitem concluir que embora a taxa de identificação não seja muito diferente para os dois testes em causa, com a normalização geométrica das imagens todos os indivíduos do conjunto de teste  $X_{\text{teste}_3}$  são correctamente autenticados, o que não se verifica com o conjunto de teste  $X_{\text{teste}_2}$ .

Estes resultados fundamentam que se efectue correcção geométrica às imagens da base de dados e às imagens sintetizadas a partir das imagens extraídas da sequência vídeo.

## 5.2.1 Redução da Dimensionalidade dos Dados

Com o intuito de avaliar a consistência do método de análise de componentes independentes aplicado ao reconhecimento de imagens, analisou-se o comportamento das taxas de identificação e de autenticação de acordo com a redução do número de componentes independentes estimadas. A redução de dimensionalidade é obtida através da redução do número de componentes principais, descartando as componentes principais com menor variância e portanto as componentes que contêm menos informação a respeito dos dados (como referido na secção 2.3.1).

Na Figura 5.7 ilustra-se a relação entre a taxa de identificação e o número de componentes independentes estimadas utilizando as metodologias com e sem normalização geométrica das imagens.

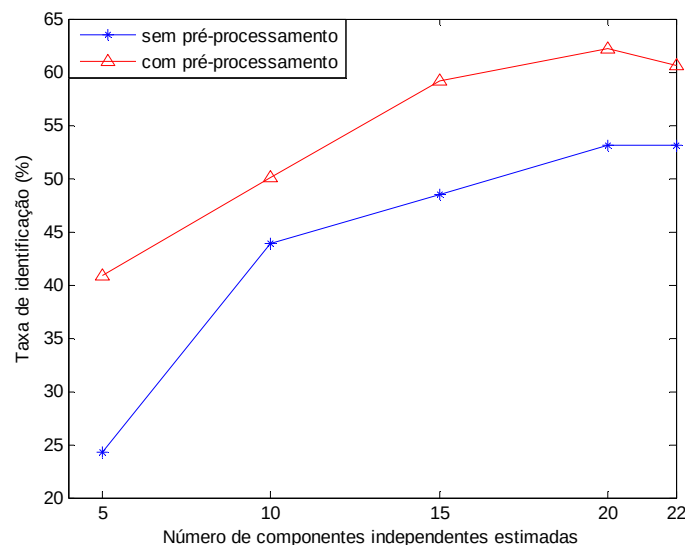


**Figura 5.7 – Relação entre a taxa de identificação e o número de componentes independentes estimadas, considerando as metodologias com e sem normalização geométrica.**

Por observação da Figura 5.7 verifica-se, em geral, que a redução do número de componentes independentes estimadas acarreta, em ambas as metodologias, um decréscimo da taxa de identificação. Verifica-se ainda que a metodologia com correcção geométrica das imagens conduz, independentemente do número de componentes estimadas, a taxas de identificação superiores. Uma análise mais detalhada permite observar que considerando entre 22 a 15 componentes existe uma diferença de cerca de

10% nas taxas de identificação. Esta diferença diminui aquando da estimação de 10 componentes, voltando a aumentar para cerca de 15% no caso de se estimarem apenas 5 componentes independentes.

O desempenho da taxa de autenticação em função do número de componentes independentes estimadas pode ser analisado na Figura 5.8, considerando também as metodologias com e sem correcção geométrica das imagens.

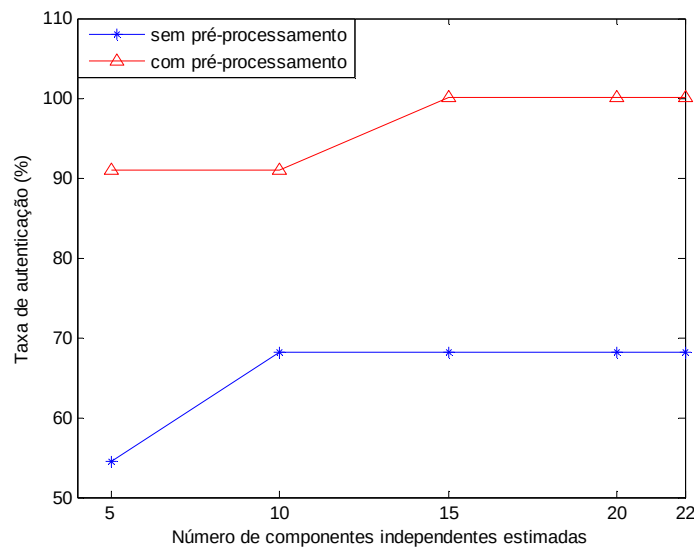


**Figura 5.8 – Relação entre a taxa de autenticação e o número de componentes independentes estimadas, considerando as metodologias com e sem normalização geométrica.**

A Figura 5.8 ilustra que utilizando a metodologia com correcção geométrica das imagens se obtêm taxas de autenticação superiores do que utilizando a metodologia sem normalização geométrica, independentemente do número de componentes estimadas.

Verifica-se ainda que estimando entre 15 a 22 componentes e utilizando a metodologia com correcção geométrica das imagens se obtêm taxas de autenticação de 100%, sendo que a metodologia sem correcção atinge o máximo da taxa de autenticação aos 68%.

## 5.3 Autenticação Utilizando as Vistas Frontais Sintetizadas Com Pré-Processamento

A escolha da utilização das vistas sintetizadas, através da técnica *View-Morphing*, a partir das imagens extraídas da sequência vídeo, e a opção de efectuar pré-processamento às mesmas, são decisões baseadas nos resultados apresentados e discutidos nas secções 5.1 e 5.2. Deste modo, fica definida a construção do sistema automático de visão, cuja principal finalidade é controlar o acesso de “clientes” a áreas restritas e evitar a entrada de “impostores”, em tempo real e a partir de sequências de vídeo com poucas restrições em relação a iluminação e movimentação dos indivíduos que requerem acesso.

A base de dados deste sistema será constituída pelas imagens pré-processadas, em pose frontal e neutra. A autenticação de “clientes” e a não autenticação/reconhecimento de “impostores” serão efectuadas por recurso às imagens sintetizadas, também estas pré-processadas.

A metodologia em que se baseia a definição do valor limiar acima do qual os indivíduos não são autenticados é comum a todos os indivíduos. Contudo, estes valores de *threshold* são dinâmicos, ou seja, cada indivíduo tem associado um valor limiar abaixo do qual é autenticado. Note-se que, deste modo, sempre que seja necessário introduzir um novo indivíduo na base de dados é preciso efectuar um teste de identificação por forma a estabelecer um valor limiar para todos os indivíduos.

Na Tabela 5.4 apresentam-se as distâncias de todas as representações ICA do conjunto de teste  $\mathbf{X}_{\text{teste}_3}$  (imagens sintetizadas após equalização do histograma e correcção geométrica) às representações do conjunto de treino  $\mathbf{X}_{\text{treino}_2}$  (imagens pré-processadas).

**Tabela 5.4 - Distâncias das representações ICA das imagens sintetizadas (com equalização do histograma e com correção geométrica) às representações do conjunto de treino  $X_{\text{treino}_2}$ .**

|                                               |    | Indivíduos constituintes do conjunto de treino |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      | Threshold    |
|-----------------------------------------------|----|------------------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|--------------|
|                                               |    | 1                                              | 2                    | 3                    | 4                    | 5                    | 6                    | 7                    | 8                    | 9                    | 10                   | 11                   | 12                   | 13                   | 14                   | 15                   | 16                   | 17                   | 18                   | 19                   | 20                   | 21                   | 22                   | $\mu-\sigma$ |
| Indivíduos constituintes do conjunto de teste | 1  | 0,26<br>0,14<br>0,18                           | 0,55<br>0,56<br>0,61 | 0,49<br>0,49<br>0,52 | 0,55<br>0,60<br>0,58 | 0,27<br>0,29<br>0,32 | 0,33<br>0,33<br>0,43 | 0,27<br>0,18<br>0,25 | 0,60<br>0,60<br>0,63 | 0,30<br>0,32<br>0,35 | 0,35<br>0,26<br>0,40 | 0,37<br>0,42<br>0,29 | 0,32<br>0,38<br>0,63 | 0,57<br>0,60<br>0,46 | 0,42<br>0,43<br>0,52 | 0,47<br>0,30<br>0,33 | 0,37<br>0,23<br>0,22 | 0,26<br>0,51<br>0,52 | 0,53<br>0,50<br>0,53 | 0,54<br>0,45<br>0,43 | 0,45<br>0,27<br>0,43 | 0,32<br>0,42<br>0,30 | 0,45<br>0,61<br>0,44 | 0,28         |
|                                               | 2  | 0,38<br>0,38<br>0,36                           | 0,43<br>0,13<br>0,15 | 0,58<br>0,56<br>0,51 | 0,50<br>0,65<br>0,66 | 0,43<br>0,58<br>0,58 | 0,15<br>0,35<br>0,31 | 0,39<br>0,48<br>0,54 | 0,52<br>0,63<br>0,69 | 0,58<br>0,54<br>0,50 | 0,39<br>0,36<br>0,38 | 0,54<br>0,63<br>0,36 | 0,28<br>0,35<br>0,98 | 0,81<br>0,96<br>0,67 | 0,60<br>0,68<br>0,57 | 0,69<br>0,65<br>0,73 | 0,64<br>0,66<br>0,50 | 0,50<br>0,88<br>0,70 | 0,84<br>0,79<br>0,86 | 0,70<br>0,56<br>0,82 | 0,45<br>0,57<br>0,53 | 0,46<br>0,61<br>0,64 | 0,49<br>0,57<br>0,59 | 0,38         |
|                                               | 3  | 0,57<br>0,52<br>0,62                           | 0,94<br>0,64<br>0,87 | 0,25<br>0,46<br>0,31 | 0,53<br>0,44<br>0,44 | 0,57<br>0,48<br>0,61 | 0,50<br>0,48<br>0,51 | 0,48<br>0,42<br>0,51 | 0,50<br>0,54<br>0,47 | 0,69<br>0,78<br>0,79 | 0,62<br>0,54<br>0,62 | 0,56<br>0,60<br>0,59 | 0,50<br>0,32<br>0,48 | 0,28<br>0,56<br>0,31 | 0,48<br>0,56<br>0,52 | 0,69<br>0,77<br>0,79 | 0,74<br>0,75<br>0,51 | 0,47<br>0,41<br>0,40 | 0,33<br>0,64<br>0,40 | 0,32<br>0,54<br>0,38 | 0,39<br>0,55<br>0,43 | 0,61<br>0,36<br>0,61 | 0,39<br>0,54<br>0,47 | 0,39         |
|                                               | 4  | 0,85<br>0,75<br>0,69                           | 0,71<br>0,66<br>0,67 | 0,84<br>0,79<br>0,94 | 0,31<br>0,32<br>0,19 | 0,61<br>0,59<br>0,47 | 0,54<br>0,47<br>0,37 | 0,68<br>0,59<br>0,54 | 0,31<br>0,33<br>0,37 | 1,01<br>0,90<br>0,83 | 0,88<br>0,76<br>0,69 | 0,66<br>0,69<br>0,66 | 0,77<br>0,67<br>0,74 | 0,72<br>0,72<br>0,85 | 0,96<br>1,08<br>1,03 | 0,98<br>0,96<br>0,84 | 0,81<br>0,65<br>0,67 | 0,89<br>0,86<br>0,98 | 0,66<br>0,64<br>0,82 | 0,86<br>0,80<br>0,91 | 0,52<br>0,44<br>0,41 | 0,63<br>0,62<br>0,67 | 0,52<br>0,62<br>0,67 | 0,50         |
|                                               | 5  | 0,61<br>0,67<br>0,54                           | 0,75<br>0,92<br>0,81 | 0,65<br>0,71<br>0,59 | 0,29<br>0,39<br>0,33 | 0,35<br>0,33<br>0,34 | 0,43<br>0,44<br>0,47 | 0,52<br>0,44<br>0,38 | 0,45<br>0,32<br>0,45 | 0,80<br>0,89<br>0,82 | 0,69<br>0,70<br>0,60 | 0,54<br>0,51<br>0,57 | 0,45<br>0,66<br>0,63 | 0,68<br>0,83<br>0,70 | 0,73<br>0,91<br>0,80 | 0,91<br>0,77<br>0,79 | 0,83<br>0,54<br>0,35 | 0,44<br>0,68<br>0,65 | 0,71<br>0,56<br>0,54 | 0,62<br>0,81<br>0,66 | 0,69<br>0,56<br>0,66 | 0,34<br>0,29<br>0,28 | 0,50<br>0,44<br>0,45 | 0,41         |
|                                               | 6  | 0,52<br>0,49<br>0,68                           | 0,80<br>0,82<br>0,89 | 0,72<br>0,64<br>0,78 | 0,34<br>0,35<br>0,27 | 0,26<br>0,26<br>0,33 | 0,22<br>0,22<br>0,36 | 0,35<br>0,36<br>0,47 | 0,29<br>0,34<br>0,27 | 0,63<br>0,62<br>0,78 | 0,60<br>0,61<br>0,75 | 0,36<br>0,37<br>0,35 | 0,67<br>0,58<br>0,78 | 0,70<br>0,66<br>0,65 | 0,82<br>0,77<br>0,88 | 0,72<br>0,69<br>0,83 | 0,61<br>0,67<br>0,63 | 0,56<br>0,49<br>0,68 | 0,66<br>0,63<br>0,65 | 0,60<br>0,56<br>0,60 | 0,60<br>0,71<br>0,87 | 0,79<br>0,25<br>0,34 | 0,23<br>0,30<br>0,39 | 0,34         |
|                                               | 7  | 0,22<br>0,23<br>0,30                           | 0,48<br>0,67<br>0,63 | 0,54<br>0,47<br>0,49 | 0,57<br>0,52<br>0,56 | 0,29<br>0,27<br>0,26 | 0,22<br>0,30<br>0,29 | 0,15<br>0,20<br>0,13 | 0,40<br>0,44<br>0,27 | 0,30<br>0,52<br>0,36 | 0,22<br>0,26<br>0,32 | 0,36<br>0,41<br>0,26 | 0,41<br>0,28<br>0,42 | 0,63<br>0,55<br>0,55 | 0,57<br>0,59<br>0,42 | 0,96<br>0,63<br>0,56 | 0,51<br>0,40<br>0,37 | 0,37<br>0,29<br>0,35 | 0,33<br>0,54<br>0,51 | 0,64<br>0,49<br>0,39 | 0,54<br>0,41<br>0,43 | 0,44<br>0,32<br>0,25 | 0,29<br>0,35<br>0,30 | 0,27         |
|                                               | 8  | 0,75<br>0,77<br>0,78                           | 0,81<br>0,84<br>0,81 | 0,84<br>0,49<br>0,91 | 0,44<br>0,43<br>0,40 | 0,41<br>0,47<br>0,38 | 0,41<br>0,54<br>0,40 | 0,55<br>0,09<br>0,64 | 0,09<br>0,07<br>0,13 | 0,79<br>0,79<br>0,77 | 0,66<br>0,68<br>0,71 | 0,42<br>0,50<br>0,41 | 0,81<br>0,60<br>0,85 | 0,61<br>0,96<br>0,71 | 0,96<br>0,95<br>0,99 | 0,96<br>0,67<br>0,95 | 0,67<br>0,82<br>0,65 | 0,80<br>0,80<br>0,86 | 0,66<br>0,60<br>0,73 | 0,80<br>0,78<br>0,85 | 0,49<br>0,47<br>0,53 | 0,49<br>0,48<br>0,52 | 0,44                 |              |
|                                               | 9  | 0,24<br>0,28<br>0,25                           | 0,52<br>0,49<br>0,57 | 0,66<br>0,74<br>0,65 | 0,73<br>0,60<br>0,57 | 0,25<br>0,29<br>0,25 | 0,39<br>0,29<br>0,37 | 0,36<br>0,36<br>0,36 | 0,60<br>0,59<br>0,57 | 0,20<br>0,24<br>0,26 | 0,35<br>0,36<br>0,33 | 0,32<br>0,33<br>0,31 | 0,51<br>0,79<br>0,61 | 0,74<br>0,68<br>0,47 | 0,53<br>0,68<br>0,47 | 0,33<br>0,45<br>0,24 | 0,25<br>0,26<br>0,40 | 0,35<br>0,49<br>0,60 | 0,57<br>0,70<br>0,60 | 0,61<br>0,73<br>0,63 | 0,59<br>0,62<br>0,51 | 0,30<br>0,39<br>0,35 | 0,42<br>0,49<br>0,43 | 0,30         |
|                                               | 10 | 0,27<br>0,29<br>0,29                           | 0,44<br>0,60<br>0,45 | 0,61<br>0,62<br>0,60 | 0,65<br>0,46<br>0,64 | 0,31<br>0,26<br>0,34 | 0,36<br>0,22<br>0,33 | 0,33<br>0,23<br>0,34 | 0,58<br>0,43<br>0,63 | 0,35<br>0,44<br>0,48 | 0,18<br>0,26<br>0,25 | 0,50<br>0,37<br>0,63 | 0,63<br>0,34<br>0,55 | 0,82<br>0,67<br>0,89 | 0,58<br>0,70<br>0,64 | 0,43<br>0,60<br>0,55 | 0,39<br>0,35<br>0,49 | 0,36<br>0,40<br>0,37 | 0,70<br>0,65<br>0,80 | 0,74<br>0,62<br>0,81 | 0,62<br>0,52<br>0,66 | 0,34<br>0,32<br>0,33 | 0,53<br>0,40<br>0,63 | 0,32         |
|                                               | 11 | 0,53<br>0,48<br>0,41                           | 0,75<br>0,72<br>0,67 | 0,54<br>0,67<br>0,68 | 0,72<br>0,56<br>0,51 | 0,28<br>0,17<br>0,17 | 0,47<br>0,36<br>0,33 | 0,42<br>0,32<br>0,30 | 0,30<br>0,23<br>0,37 | 0,42<br>0,45<br>0,46 | 0,60<br>0,50<br>0,47 | 0,09<br>0,12<br>0,23 | 0,65<br>0,63<br>0,53 | 0,42<br>0,50<br>0,54 | 0,58<br>0,71<br>0,71 | 0,51<br>0,59<br>0,62 | 0,25<br>0,32<br>0,33 | 0,61<br>0,53<br>0,48 | 0,36<br>0,54<br>0,60 | 0,37<br>0,46<br>0,55 | 0,57<br>0,66<br>0,64 | 0,42<br>0,66<br>0,19 | 0,24<br>0,27<br>0,35 | 0,29         |
|                                               | 12 | 0,48<br>0,48<br>0,40                           | 0,83<br>0,79<br>0,72 | 0,31<br>0,19<br>0,22 | 0,62<br>0,67<br>0,63 | 0,54<br>0,49<br>0,48 | 0,50<br>0,52<br>0,57 | 0,32<br>0,34<br>0,35 | 0,45<br>0,51<br>0,49 | 0,65<br>0,59<br>0,62 | 0,43<br>0,48<br>0,41 | 0,56<br>0,46<br>0,52 | 0,29<br>0,26<br>0,32 | 0,24<br>0,46<br>0,37 | 0,47<br>0,66<br>0,72 | 0,77<br>0,66<br>0,66 | 0,68<br>0,30<br>0,25 | 0,30<br>0,44<br>0,36 | 0,44<br>0,30<br>0,36 | 0,33<br>0,33<br>0,27 | 0,51<br>0,51<br>0,27 | 0,45<br>0,48<br>0,51 | 0,36<br>0,46<br>0,36 | 0,30         |
|                                               | 13 | 0,67<br>0,74<br>0,57                           | 1,04<br>1,11<br>0,98 | 0,30<br>0,33<br>0,25 | 0,77<br>0,80<br>0,79 | 0,67<br>0,74<br>0,66 | 0,73<br>0,80<br>0,69 | 0,49<br>0,54<br>0,41 | 0,48<br>0,51<br>0,52 | 0,72<br>0,76<br>0,64 | 0,60<br>0,70<br>0,51 | 0,52<br>0,55<br>0,48 | 0,58<br>0,62<br>0,43 | 0,07<br>0,08<br>0,08 | 0,45<br>0,47<br>0,39 | 0,76<br>0,78<br>0,70 | 0,57<br>0,63<br>0,51 | 0,59<br>0,63<br>0,50 | 0,27<br>0,25<br>0,27 | 0,26<br>0,24<br>0,24 | 0,37<br>0,40<br>0,30 | 0,76<br>0,78<br>0,69 | 0,46<br>0,48<br>0,44 | 0,33         |
|                                               | 14 | 0,34<br>0,35<br>0,33                           | 0,64<br>0,76<br>0,74 | 0,14<br>0,15<br>0,14 | 0,95<br>0,86<br>0,81 | 0,79<br>0,70<br>0,61 | 0,78<br>0,71<br>0,62 | 0,47<br>0,38<br>0,33 | 0,76<br>0,72<br>0,59 | 0,40<br>0,47<br>0,42 | 0,32<br>0,35<br>0,33 | 0,64<br>0,59<br>0,47 | 0,38<br>0,47<br>0,44 | 0,37<br>0,45<br>0,20 | 0,11<br>0,24<br>0,21 | 0,43<br>0,50<br>0,49 | 0,45<br>0,37<br>0,36 | 0,31<br>0,39<br>0,37 | 0,23<br>0,19<br>0,18 | 0,32<br>0,31<br>0,24 | 0,14<br>0,21<br>0,18 | 0,81<br>0,77<br>0,69 | 0,53<br>0,50<br>0,41 | 0,23         |
|                                               | 15 | 0,37<br>0,30<br>0,26                           | 0,64<br>0,64<br>0,61 | 0,65<br>0,31<br>0,30 | 0,81<br>0,54<br>0,95 | 0,38<br>0,62<br>0,51 | 0,61<br>0,41<br>0,60 | 0,45<br>0,71<br>0,60 | 0,71<br>0,22<br>0,36 | 0,49<br>0,71<br>0,63 | 0,26<br>0,32<br>0,28 | 0,49<br>0,42<br>0,40 | 0,31<br>0,56<br>0,51 | 0,54<br>0,49<br>0,45 | 0,66<br>0,32<br>0,33 | 0,47<br>0,08<br>0,15 | 0,28<br>0,22<br>0,18 | 0,45<br>0,41<br>0,39 | 0,50<br>0,42<br>0,24 | 0,60<br>0,57<br>0,39 | 0,67<br>0,42<br>0,36 | 0,53<br>0,65<br>0,62 | 0,42<br>0,43<br>0,44 | 0,27         |
|                                               | 16 | 0,26<br>0,42<br>0,26                           | 0,63<br>0,77<br>0,61 | 0,60<br>0,53<br>0,51 | 0,85<br>0,71<br>0,64 | 0,35<br>0,28<br>0,25 | 0,51<br>0,50<br>0,32 | 0,39<br>0,32<br>0,26 | 0,60<br>0,26<br>0,37 | 0,15<br>0,37<br>0,24 | 0,33<br>0,45<br>0,27 | 0,31<br>0,12<br>0,49 | 0,63<br>0,63<br>0,44 | 0,61<br>0,30<br>0,52 | 0,48<br>0,54<br>0,45 | 0,24<br>0,18<br>0,45 | 0,10<br>0,50<br>0,15 | 0,46<br>0,50<br>0,41 | 0,45<br>0,34<br>0,46 | 0,57<br>0,31<br>0,44 | 0,51<br>0,46<br>0,42 | 0,48<br>0,44<br>0,37 | 0,45<br>0,23<br>0,28 | 0,26         |
|                                               | 17 | 0,35<br>0,35<br>0,40                           | 0,66<br>0,75<br>0,69 | 0,41<br>0,32<br>0,38 | 0,54<br>0,58<br>0,60 | 0,39<br>0,41<br>0,58 | 0,52<br>0,30<br>0,41 | 0,33<br>0,30<br>0,75 | 0,67<br>0,65<br>0,66 | 0,58<br>0,52<br>0,49 | 0,46<br>0,45<br>0,66 | 0,55<br>0,55<br>0,69 | 0,27<br>0,35<br>0,21 | 0,56<br>0,45<br>0,57 | 0,44<br>0,38<br>0,44 | 0,56<br>0,52<br>0,67 | 0,59<br>0,62<br>0,78 | 0,18<br>0,16<br>0,21 | 0,46<br>0,37<br>0,52 | 0,47<br>0,38<br>0,47 | 0,47<br>0,42<br>0,41 | 0,37<br>0,34<br>0,48 | 0,49<br>0,44<br>0,59 | 0,34         |
|                                               | 18 | 0,43<br>0,36<br>0,41                           | 0,83<br>0,82<br>0,62 | 0,18<br>0,23<br>0,48 | 0,59<br>0,70<br>0,61 | 0,49<br>0,44<br>0,32 | 0,71<br>0,67<br>0,57 | 0,36<br>0,34<br>0,39 | 0,54<br>0,46<br>0,64 | 0,48<br>0,54<br>0,50 | 0,49<br>0,42<br>0,49 | 0,37<br>0,42<br>0,41 | 0,48<br>0,22<br>0,33 | 0,19<br>0,34<br>0,59 | 0,27<br>0,42<br>0,43 | 0,39<br>0,52<br>0,46 | 0,40<br>0,41<br>0,46 | 0,36<br>0,32<br>0,32 | 0,09<br>0,11<br>0,45 | 0,18<br>0,23<br>0,51 | 0,29<br>0,29<br>0,51 | 0,53<br>0,57<br>0,29 | 0,31<br>0,31<br>0,44 | 0,27         |
|                                               | 19 | 0,47<br>0,39<br>0,44                           | 0,92<br>0,80<br>0,85 | 0,24<br>0,20<br>0,42 | 0,59<br>0,73<br>0,45 | 0,49<br>0,42<br>0,27 | 0,55<br>0,54<br>0,31 | 0,37<br>0,25<br>0,25 | 0,54<br>0,42<br>0,58 | 0,65<br>0,46<br>0,44 | 0,52<br>0,41<br>0,44 | 0,46<br>0,31<br>0,37 | 0,42<br>0,44<br>0,47 | 0,19<br>0,16<br>0,34 | 0,42<br>0,32<br>0,58 | 0,65<br>0,49<br>0,62 | 0,54<br>0,38<br>0,48 | 0,41<br>0,35<br>0,42 | 0,23<br>0,16<br>0,40 | 0,24<br>0,13<br>0,40 | 0,31<br>0,27<br>0,36 | 0,57<br>0,47<br>0,30 | 0,37<br>0,25<br>0,33 | 0,26         |
|                                               | 20 | 0,35<br>0,29<br>0,34                           | 0,78<br>0,73<br>0,71 | 0,12<br>0,14<br>0,12 | 0,79<br>0,72<br>0,85 | 0,52<br>0,53<br>0,65 | 0,61<br>0,59<br>0,69 | 0,34<br>0,34<br>0,36 | 0,55<br>0,53<br>0,61 | 0,45<br>0,49<br>0,42 | 0,39<br>0,36<br>0,36 | 0,55<br>0,50<br>0,51 | 0,43<br>0,37<br>0,35 | 0,39<br>0,26<br>0,21 | 0,18<br>0,25<br>0,18 | 0,42<br>0,53<br>0,48 | 0,41<br>0,46<br>0,42 | 0,32<br>0,28<br>0,30 | 0,16<br>0,18<br>0,15 | 0,22<br>0,26<br>0,20 | 0,19<br>0,18<br>0,12 | 0,59<br>0,63<br>0,68 | 0,37<br>0,37<br>0,42 | 0,22         |
|                                               | 21 | 0,62<br>0,52<br>0,46                           | 0,86<br>0,84<br>0,82 | 0,53<br>0,62<br>0,57 | 0,45<br>0,47<br>0,51 | 0,41<br>0,33<br>0,31 | 0,56<br>0,52<br>0,47 | 0,48<br>0,37<br>0,31 | 0,42<br>0,37<br>0,57 | 0,78<br>0,73<br>0,71 | 0,68<br>0,60<br>0,57 | 0,50<br>0,55<br>0,50 | 0,44<br>0,47<br>0,38 | 0,52<br>0,70<br>0,65 | 0,60<br>0,62<br>0,66 | 0,81<br>0,67<br>0,65 | 0,79<br>0,68<br>0,63 | 0,39<br>0,32<br>0,34 | 0,50<br>0,63<br>0,59 | 0,42<br>0,60<br>0,53 | 0,59<br>0,60<br>0,62 | 0,29<br>0,22<br>0,20 | 0,39<br>0,48<br>0,43 | 0,39         |
|                                               | 22 | 0,52<br>0,57<br>0,47                           | 0,84<br>0,99<br>0,64 | 0,62<br>0,41<br>0,55 | 0,47<br>0,45<br>0,46 | 0,33<br>0,45<br>0,19 | 0,52<br>0,49<br>0,32 | 0,37<br>0,49<br>0,33 | 0,61<br>0,57<br>0,33 | 0,73<br>0,74<br>0,32 | 0,60<br>0,65<br>0,32 | 0,55<br>0,54<br>0,32 | 0,47<br>0,54<br>0,57 | 0,70<br>0,48<br>0,38 | 0,62<br>0,31<br>0,49 | 0,67<br>0,57<br>0,67 | 0,68<br>0,68<br>0,50 | 0,32<br>0,46<br>0,44 | 0,33<br>0,37<br>0,52 | 0,32<br>0,35<br>0,49 | 0,46<br>0,44<br>0,55 | 0,60<br>0,42<br>0,22 | 0,29<br>0,42<br>0,20 | 0,32         |

Por observação desta tabela pode-se verificar que, construindo o sistema de autenticação de clientes como referido no início desta secção, todos os indivíduos presentes na base de dados são correctamente autenticados, permitindo assim o acesso destes a áreas restritas.

## 5.4 Testes de Rejeição de Impostores

É patente que o impedimento de entrada de “impostores” em áreas de acesso restrito é uma questão que assume relevo superior à permissão da entrada de “clientes” autorizados. Por esta razão, e para avaliar a fiabilidade e a segurança com que este sistema automático rejeita a entrada de impostores, realizaram-se testes de rejeição de impostores, do tipo “*Leave-One-Out*”, que consistem em construir um conjunto de treino com  $n-1$  indivíduos (sendo  $n$  o número total de indivíduos presentes na base de dados, neste caso 22 indivíduos) e avaliar as disparidades e semelhanças entre as representações ICA deste conjunto e as representações do conjunto de teste formado pelas imagens sintetizadas do indivíduo “deixado de fora”. Este teste caracteriza-se por alterar o indivíduo que é “deixado de fora”, de forma que todos os indivíduos sejam testados a partir do treinamento dos restantes.

Considerando-se que se deixa de fora o indivíduo  $j$  (com  $j \in \{1, 2, \dots, 22\}$ ), constrói-se um conjunto de treino com os restantes indivíduos e posteriormente comparam-se as representações ICA das imagens sintetizadas relativas ao indivíduo  $j$  com todas as representações do conjunto de treino. Estas distâncias permitem definir um valor de *threshold* dado pela (Eq. 5.67):

$$T_i = \mu_i - 2\sigma_i \quad (\text{Eq. 5.67})$$

sendo  $\mu_i$  e  $\sigma_i$  dados pela (Eq. 5.65) e com  $N_{\text{images}} = 3$  e  $N_{\text{ind}} = 21$ .

Deste modo, os impostores são autenticados caso as distâncias de, pelo menos, duas das três imagens sintetizadas a uma representação do conjunto de treino sejam inferiores ao valor limiar calculado.

Na Tabela 5.5 apresentam-se as distâncias das representações ICA das imagens sintetizadas e pré-processadas do indivíduo  $j$  a todas as representações do conjunto de treino compostos pelos restantes indivíduos, percorrendo  $j$  de 1 a 22.

**Tabela 5.5 - Distâncias das representações ICA das imagens sintetizadas (com pré-processamento) do indivíduo j (j=1,2,...,22) às representações do conjunto de treino constituído pelos restantes 21 elementos.**

|                             |    | Indivíduos constituintes do conjunto de treino |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      |                      | Threshold     |
|-----------------------------|----|------------------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|---------------|
|                             |    | 1                                              | 2                    | 3                    | 4                    | 5                    | 6                    | 7                    | 8                    | 9                    | 10                   | 11                   | 12                   | 13                   | 14                   | 15                   | 16                   | 17                   | 18                   | 19                   | 20                   | 21                   | 22                   | $\mu-2\sigma$ |
| Indivíduo "deixado de fora" | 1  |                                                | 0,55<br>0,56<br>0,60 | 0,49<br>0,49<br>0,51 | 0,55<br>0,60<br>0,57 | 0,27<br>0,29<br>0,29 | 0,33<br>0,32<br>0,42 | 0,27<br>0,17<br>0,23 | 0,60<br>0,60<br>0,63 | 0,30<br>0,31<br>0,36 | 0,35<br>0,25<br>0,33 | 0,37<br>0,42<br>0,39 | 0,32<br>0,38<br>0,26 | 0,57<br>0,60<br>0,62 | 0,42<br>0,44<br>0,45 | 0,47<br>0,43<br>0,50 | 0,37<br>0,29<br>0,31 | 0,26<br>0,22<br>0,20 | 0,53<br>0,50<br>0,51 | 0,54<br>0,49<br>0,51 | 0,45<br>0,44<br>0,41 | 0,32<br>0,26<br>0,28 | 0,45<br>0,42<br>0,43 | 0,17          |
|                             | 2  | 0,37<br>0,32<br>0,31                           |                      | 0,57<br>0,49<br>0,46 | 0,49<br>0,60<br>0,62 | 0,43<br>0,52<br>0,53 | 0,14<br>0,28<br>0,26 | 0,39<br>0,42<br>0,50 | 0,52<br>0,58<br>0,65 | 0,39<br>0,48<br>0,46 | 0,53<br>0,29<br>0,34 | 0,27<br>0,57<br>0,60 | 0,32<br>0,28<br>0,31 | 0,80<br>0,93<br>0,96 | 0,63<br>0,59<br>0,63 | 0,68<br>0,59<br>0,70 | 0,64<br>0,49<br>0,45 | 0,49<br>0,83<br>0,83 | 0,69<br>0,74<br>0,78 | 0,55<br>0,53<br>0,49 | 0,45<br>0,51<br>0,53 | 0,49<br>0,55<br>0,55 | 0,19                 |               |
|                             | 3  | 0,56<br>0,52<br>0,61                           | 0,93<br>0,64<br>0,86 |                      | 0,52<br>0,44<br>0,44 | 0,56<br>0,48<br>0,60 | 0,49<br>0,47<br>0,50 | 0,47<br>0,41<br>0,50 | 0,49<br>0,53<br>0,47 | 0,68<br>0,77<br>0,78 | 0,61<br>0,60<br>0,59 | 0,54<br>0,31<br>0,47 | 0,49<br>0,55<br>0,30 | 0,27<br>0,56<br>0,51 | 0,68<br>0,77<br>0,78 | 0,73<br>0,74<br>0,78 | 0,46<br>0,41<br>0,50 | 0,32<br>0,64<br>0,39 | 0,31<br>0,54<br>0,38 | 0,38<br>0,55<br>0,42 | 0,60<br>0,36<br>0,60 | 0,38<br>0,53<br>0,46 | 0,25                 |               |
|                             | 4  | 0,78<br>0,70<br>0,61                           | 0,65<br>0,62<br>0,60 | 0,75<br>0,72<br>0,83 |                      | 0,54<br>0,53<br>0,39 | 0,47<br>0,43<br>0,29 | 0,61<br>0,54<br>0,46 | 0,24<br>0,29<br>0,75 | 0,94<br>0,84<br>0,62 | 0,81<br>0,71<br>0,57 | 0,58<br>0,63<br>0,65 | 0,69<br>0,61<br>0,76 | 0,64<br>0,66<br>0,90 | 0,88<br>0,79<br>0,90 | 1,08<br>1,02<br>0,94 | 0,90<br>0,90<br>0,75 | 0,75<br>0,61<br>0,59 | 0,80<br>0,80<br>0,88 | 0,60<br>0,59<br>0,74 | 0,78<br>0,74<br>0,82 | 0,44<br>0,39<br>0,32 | 0,57<br>0,57<br>0,59 | 0,29          |
|                             | 5  | 0,60<br>0,67<br>0,54                           | 0,75<br>0,92<br>0,81 | 0,65<br>0,71<br>0,58 | 0,28<br>0,29<br>0,39 |                      | 0,43<br>0,44<br>0,47 | 0,52<br>0,44<br>0,38 | 0,44<br>0,32<br>0,44 | 0,80<br>0,89<br>0,82 | 0,68<br>0,70<br>0,60 | 0,53<br>0,51<br>0,66 | 0,44<br>0,66<br>0,63 | 0,68<br>0,73<br>0,83 | 0,73<br>0,91<br>0,91 | 0,91<br>0,77<br>0,80 | 0,83<br>0,91<br>0,79 | 0,43<br>0,54<br>0,34 | 0,70<br>0,68<br>0,64 | 0,61<br>0,56<br>0,53 | 0,68<br>0,81<br>0,66 | 0,34<br>0,28<br>0,27 | 0,49<br>0,44<br>0,44 | 0,24          |
|                             | 6  | 0,51<br>0,48<br>0,68                           | 0,79<br>0,82<br>0,89 | 0,72<br>0,63<br>0,78 | 0,32<br>0,33<br>0,26 | 0,23<br>0,23<br>0,32 |                      | 0,33<br>0,34<br>0,46 | 0,27<br>0,32<br>0,26 | 0,62<br>0,61<br>0,78 | 0,59<br>0,60<br>0,75 | 0,33<br>0,35<br>0,34 | 0,66<br>0,57<br>0,78 | 0,69<br>0,66<br>0,65 | 0,82<br>0,77<br>0,88 | 0,71<br>0,69<br>0,83 | 0,60<br>0,66<br>0,63 | 0,55<br>0,48<br>0,68 | 0,65<br>0,63<br>0,64 | 0,58<br>0,55<br>0,60 | 0,78<br>0,71<br>0,87 | 0,20<br>0,22<br>0,33 | 0,32<br>0,28<br>0,38 | 0,15          |
|                             | 7  | 0,20<br>0,23<br>0,29                           | 0,47<br>0,67<br>0,63 | 0,54<br>0,47<br>0,49 | 0,57<br>0,52<br>0,56 | 0,28<br>0,27<br>0,25 | 0,21<br>0,30<br>0,27 |                      | 0,39<br>0,44<br>0,26 | 0,29<br>0,52<br>0,35 | 0,20<br>0,26<br>0,31 | 0,35<br>0,41<br>0,24 | 0,40<br>0,28<br>0,47 | 0,63<br>0,55<br>0,55 | 0,57<br>0,59<br>0,56 | 0,50<br>0,63<br>0,36 | 0,36<br>0,40<br>0,34 | 0,31<br>0,29<br>0,51 | 0,64<br>0,54<br>0,38 | 0,53<br>0,49<br>0,38 | 0,43<br>0,41<br>0,42 | 0,28<br>0,32<br>0,24 | 0,35<br>0,35<br>0,29 | 0,15          |
|                             | 8  | 0,68<br>0,70<br>0,72                           | 0,75<br>0,79<br>0,76 | 0,78<br>0,77<br>0,85 | 0,34<br>0,38<br>0,30 | 0,31<br>0,32<br>0,28 | 0,31<br>0,36<br>0,30 | 0,47<br>0,44<br>0,57 |                      | 0,73<br>0,72<br>0,72 | 0,60<br>0,60<br>0,65 | 0,32<br>0,30<br>0,31 | 0,75<br>0,73<br>0,80 | 0,54<br>0,51<br>0,64 | 0,92<br>0,91<br>0,95 | 0,91<br>0,90<br>0,91 | 0,60<br>0,59<br>0,58 | 0,76<br>0,74<br>0,82 | 0,78<br>0,73<br>0,81 | 0,59<br>0,51<br>0,66 | 0,74<br>0,71<br>0,80 | 0,40<br>0,35<br>0,44 | 0,40<br>0,38<br>0,44 | 0,21          |
|                             | 9  | 0,24<br>0,27<br>0,25                           | 0,51<br>0,49<br>0,57 | 0,66<br>0,75<br>0,65 | 0,73<br>0,60<br>0,57 | 0,25<br>0,28<br>0,25 | 0,39<br>0,28<br>0,37 | 0,36<br>0,36<br>0,57 | 0,60<br>0,59<br>0,57 |                      | 0,35<br>0,35<br>0,36 | 0,31<br>0,33<br>0,31 | 0,51<br>0,54<br>0,41 | 0,74<br>0,79<br>0,63 | 0,53<br>0,68<br>0,61 | 0,32<br>0,44<br>0,47 | 0,25<br>0,25<br>0,24 | 0,34<br>0,49<br>0,40 | 0,57<br>0,71<br>0,60 | 0,61<br>0,73<br>0,63 | 0,59<br>0,61<br>0,51 | 0,30<br>0,39<br>0,35 | 0,42<br>0,48<br>0,43 | 0,15          |
|                             | 10 | 0,25<br>0,28<br>0,28                           | 0,42<br>0,60<br>0,44 | 0,61<br>0,62<br>0,60 | 0,65<br>0,46<br>0,64 | 0,29<br>0,26<br>0,33 | 0,34<br>0,23<br>0,33 | 0,31<br>0,43<br>0,34 | 0,57<br>0,44<br>0,63 | 0,34<br>0,44<br>0,48 |                      | 0,36<br>0,33<br>0,55 | 0,62<br>0,67<br>0,89 | 0,83<br>0,70<br>0,64 | 0,57<br>0,60<br>0,55 | 0,42<br>0,35<br>0,49 | 0,37<br>0,39<br>0,37 | 0,34<br>0,63<br>0,80 | 0,70<br>0,63<br>0,81 | 0,74<br>0,52<br>0,66 | 0,62<br>0,31<br>0,66 | 0,32<br>0,40<br>0,33 | 0,52<br>0,31<br>0,63 | 0,16          |
|                             | 11 | 0,51<br>0,47<br>0,41                           | 0,74<br>0,71<br>0,66 | 0,53<br>0,66<br>0,67 | 0,71<br>0,55<br>0,51 | 0,26<br>0,15<br>0,16 | 0,46<br>0,34<br>0,32 | 0,41<br>0,30<br>0,29 | 0,28<br>0,21<br>0,36 | 0,40<br>0,43<br>0,45 | 0,59<br>0,49<br>0,46 |                      | 0,64<br>0,62<br>0,52 | 0,40<br>0,49<br>0,53 | 0,57<br>0,70<br>0,70 | 0,50<br>0,57<br>0,62 | 0,23<br>0,30<br>0,32 | 0,60<br>0,52<br>0,48 | 0,34<br>0,53<br>0,60 | 0,35<br>0,44<br>0,54 | 0,56<br>0,65<br>0,64 | 0,40<br>0,22<br>0,18 | 0,22<br>0,25<br>0,35 | 0,15          |
|                             | 12 | 0,47<br>0,46<br>0,39                           | 0,81<br>0,78<br>0,71 | 0,29<br>0,18<br>0,20 | 0,60<br>0,65<br>0,61 | 0,52<br>0,47<br>0,55 | 0,49<br>0,51<br>0,34 | 0,31<br>0,33<br>0,48 | 0,43<br>0,50<br>0,61 | 0,63<br>0,57<br>0,39 | 0,42<br>0,44<br>0,50 | 0,54<br>0,44<br>0,31 |                      | 0,23<br>0,22<br>0,31 | 0,46<br>0,33<br>0,35 | 0,75<br>0,65<br>0,70 | 0,67<br>0,58<br>0,64 | 0,29<br>0,30<br>0,24 | 0,43<br>0,29<br>0,34 | 0,29<br>0,24<br>0,31 | 0,43<br>0,28<br>0,25 | 0,49<br>0,46<br>0,49 | 0,43<br>0,34<br>0,35 | 0,14          |
|                             | 13 | 0,64<br>0,72<br>0,54                           | 1,02<br>1,10<br>0,96 | 0,25<br>0,30<br>0,21 | 0,74<br>0,78<br>0,76 | 0,65<br>0,72<br>0,63 | 0,70<br>0,78<br>0,66 | 0,46<br>0,52<br>0,38 | 0,45<br>0,49<br>0,49 | 0,69<br>0,74<br>0,62 | 0,57<br>0,68<br>0,48 | 0,48<br>0,52<br>0,45 | 0,55<br>0,60<br>0,39 |                      | 0,42<br>0,44<br>0,36 | 0,73<br>0,75<br>0,68 | 0,54<br>0,60<br>0,48 | 0,57<br>0,61<br>0,47 | 0,23<br>0,22<br>0,23 | 0,22<br>0,22<br>0,20 | 0,33<br>0,37<br>0,26 | 0,73<br>0,75<br>0,66 | 0,43<br>0,46<br>0,41 | 0,13          |
|                             | 14 | 0,33<br>0,34<br>0,32                           | 0,63<br>0,75<br>0,73 | 0,12<br>0,14<br>0,13 | 0,93<br>0,85<br>0,80 | 0,78<br>0,69<br>0,60 | 0,76<br>0,70<br>0,61 | 0,45<br>0,37<br>0,32 | 0,74<br>0,71<br>0,58 | 0,38<br>0,46<br>0,41 | 0,31<br>0,35<br>0,32 | 0,63<br>0,58<br>0,46 | 0,36<br>0,44<br>0,43 | 0,36<br>0,23<br>0,19 |                      | 0,42<br>0,49<br>0,48 | 0,43<br>0,36<br>0,35 | 0,29<br>0,38<br>0,36 | 0,22<br>0,18<br>0,17 | 0,31<br>0,30<br>0,24 | 0,13<br>0,20<br>0,17 | 0,79<br>0,75<br>0,68 | 0,52<br>0,49<br>0,40 | 0,03          |
|                             | 15 | 0,35<br>0,25<br>0,23                           | 0,61<br>0,59<br>0,58 | 0,60<br>0,25<br>0,25 | 0,77<br>0,94<br>0,90 | 0,35<br>0,49<br>0,47 | 0,58<br>0,56<br>0,56 | 0,42<br>0,35<br>0,32 | 0,67<br>0,65<br>0,59 | 0,24<br>0,17<br>0,24 | 0,46<br>0,36<br>0,25 | 0,28<br>0,51<br>0,36 | 0,50<br>0,43<br>0,46 | 0,62<br>0,26<br>0,42 | 0,43<br>0,26<br>0,29 |                      | 0,18<br>0,17<br>0,14 | 0,42<br>0,36<br>0,36 | 0,46<br>0,23<br>0,20 | 0,57<br>0,37<br>0,35 | 0,54<br>0,31<br>0,29 | 0,39<br>0,59<br>0,58 | 0,50<br>0,38<br>0,40 | 0,08          |
|                             | 16 | 0,25<br>0,41<br>0,25                           | 0,62<br>0,77<br>0,59 | 0,59<br>0,51<br>0,50 | 0,84<br>0,70<br>0,63 | 0,34<br>0,26<br>0,23 | 0,50<br>0,49<br>0,31 | 0,37<br>0,31<br>0,24 | 0,59<br>0,25<br>0,35 | 0,14<br>0,44<br>0,25 | 0,31<br>0,11<br>0,19 | 0,30<br>0,62<br>0,29 | 0,62<br>0,51<br>0,53 | 0,60<br>0,29<br>0,44 | 0,47<br>0,51<br>0,49 | 0,22<br>0,53<br>0,44 |                      | 0,45<br>0,49<br>0,40 | 0,44<br>0,33<br>0,44 | 0,56<br>0,30<br>0,43 | 0,50<br>0,44<br>0,40 | 0,46<br>0,43<br>0,35 | 0,44<br>0,22<br>0,27 | 0,11          |
|                             | 17 | 0,34<br>0,33<br>0,39                           | 0,67<br>0,76<br>0,69 | 0,41<br>0,31<br>0,38 | 0,54<br>0,59<br>0,60 | 0,38<br>0,40<br>0,59 | 0,52<br>0,50<br>0,59 | 0,32<br>0,28<br>0,41 | 0,67<br>0,65<br>0,76 | 0,45<br>0,51<br>0,49 | 0,55<br>0,44<br>0,70 | 0,25<br>0,33<br>0,20 | 0,56<br>0,33<br>0,58 | 0,43<br>0,37<br>0,44 | 0,56<br>0,52<br>0,68 | 0,59<br>0,62<br>0,79 |                      | 0,46<br>0,37<br>0,53 | 0,46<br>0,37<br>0,47 | 0,47<br>0,41<br>0,41 | 0,36<br>0,33<br>0,48 | 0,49<br>0,43<br>0,59 | 0,23                 |               |
|                             | 18 | 0,42<br>0,34<br>0,41                           | 0,82<br>0,81<br>0,62 | 0,17<br>0,20<br>0,48 | 0,83<br>0,69<br>0,61 | 0,47<br>0,42<br>0,32 | 0,71<br>0,65<br>0,57 | 0,36<br>0,31<br>0,39 | 0,53<br>0,44<br>0,64 | 0,48<br>0,52<br>0,50 | 0,37<br>0,40<br>0,49 | 0,48<br>0,46<br>0,41 | 0,37<br>0,40<br>0,33 | 0,18<br>0,20<br>0,59 | 0,27<br>0,31<br>0,43 | 0,38<br>0,50<br>0,46 | 0,39<br>0,39<br>0,46 | 0,35<br>0,30<br>0,32 |                      | 0,17<br>0,21<br>0,51 | 0,29<br>0,27<br>0,51 | 0,53<br>0,55<br>0,29 | 0,31<br>0,29<br>0,44 | 0,13          |
|                             | 19 | 0,47<br>0,38<br>0,44                           | 0,92<br>0,80<br>0,85 | 0,24<br>0,19<br>0,42 | 0,59<br>0,74<br>0,45 | 0,49<br>0,41<br>0,27 | 0,55<br>0,53<br>0,31 | 0,37<br>0,24<br>0,25 | 0,54<br>0,41<br>0,30 | 0,65<br>0,45<br>0,58 | 0,52<br>0,40<br>0,44 | 0,46<br>0,30<br>0,43 | 0,42<br>0,14<br>0,34 | 0,19<br>0,31<br>0,58 | 0,42<br>0,49<br>0,62 | 0,65<br>0,37<br>0,48 | 0,54<br>0,34<br>0,42 | 0,40<br>0,15<br>0,40 | 0,22<br>0,34<br>0,42 |                      | 0,31<br>0,26<br>0,47 | 0,57<br>0,47<br>0,30 | 0,37<br>0,24<br>0,33 | 0,11          |
|                             | 20 | 0,35<br>0,29<br>0,33                           | 0,78<br>0,73<br>0,71 | 0,11<br>0,13<br>0,12 | 0,79<br>0,72<br>0,84 | 0,52<br>0,58<br>0,64 | 0,60<br>0,34<br>0,68 | 0,34<br>0,53<br>0,36 | 0,55<br>0,49<br>0,60 | 0,44<br>0,36<br>0,42 | 0,39<br>0,43<br>0,36 | 0,43<br>0,49<br>0,51 | 0,37<br>0,36<br>0,35 | 0,17<br>0,24<br>0,21 | 0,42<br>0,41<br>0,47 | 0,42<br>0,53<br>0,47 | 0,44<br>0,46<br>0,41 | 0,32<br>0,28<br>0,30 | 0,16<br>0,18<br>0,15 | 0,22<br>0,26<br>0,19 |                      | 0,59<br>0,62<br>0,68 | 0,37<br>0,37<br>0,42 | 0,05          |
|                             | 21 | 0,58<br>0,48<br>0,42                           | 0,82<br>0,80<br>0,78 | 0,48<br>0,56<br>0,51 | 0,41<br>0,43<br>0,46 | 0,37<br>0,29<br>0,27 | 0,52<br>0,48<br>0,43 | 0,44<br>0,33<br>0,27 | 0,38<br>0,57<br>0,52 | 0,74<br>0,68<br>0,67 | 0,64<br>0,56<br>0,53 | 0,45<br>0,51<br>0,46 | 0,40<br>0,42<br>0,34 | 0,48<br>0,65<br>0,60 | 0,55<br>0,63<br>0,61 | 0,76<br>0,64<br>0,60 | 0,74<br>0,64<br>0,58 | 0,36<br>0,29<br>0,31 | 0,45<br>0,58<br>0,54 | 0,39<br>0,56<br>0,49 | 0,54<br>0,66<br>0,57 |                      | 0,35<br>0,44<br>0,39 | 0,24          |
|                             | 22 | 0,52<br>0,57<br>0,47                           | 0,98<br>0,99<br>0,65 | 0,42<br>0,41<br>0,55 | 0,53<br>0,45<br>0,46 | 0,32<br>0,45<br>0,18 | 0,49<br>0,49<br>0,31 | 0,32<br>0,37<br>0,33 | 0,42<br>0,48<br>0,31 | 0,52<br>0,74<br>0,57 | 0,66<br>0,65<br>0,49 | 0,59<br>0,54<br>0,24 | 0,46<br>0,48<br>0,38 | 0,34<br>0,31<br>0,49 | 0,53<br>0,57<br>0,66 | 0,62<br>0,76<br>0,68 | 0,60<br>0,69<br>0,50 | 0,37<br>0,46<br>0,44 | 0,33<br>0,37<br>0,52 | 0,32<br>0,35<br>0,49 | 0,46<br>0,44<br>0,55 | 0,30<br>0,42<br>0,22 |                      | 0,17          |

A análise destes testes de rejeição de impostores permite concluir que, considerando o limiar definido no início desta secção pela (Eq. 5.67), nenhum impostor consegue aceder à área restrita. Estes resultados são válidos considerando cada um dos 22 indivíduos como impostores e construindo conjuntos de treino com os restantes 21.

Além destes resultados, verifica-se ainda que considerando estes conjuntos de treino e o valor de *threshold* para a autenticação definido pela (Eq. 5.64) todos os clientes (21 indivíduos constituintes do conjunto de treino) são correctamente autenticados.

Note-se que os resultados obtidos e analisados neste capítulo são válidos apenas e somente para esta base de dados e para as sequências vídeo captadas. Não é possível extrapolar que o sistema construído e os valores de *threshold* definidos serão funcionais e viáveis para quaisquer indivíduos e quaisquer sequências vídeo.

# **6 CONCLUSÕES E PERSPECTIVAS FUTURAS**

## **6.1 Conclusões**

A presente dissertação de mestrado permite exemplificar a aplicabilidade da análise de componentes independentes ao reconhecimento de pessoas por recurso à análise de imagens, nomeadamente na autenticação de “clientes”.

A revisão bibliográfica efectuada e apresentada no segundo capítulo relativa a métodos passíveis de utilização no reconhecimento de pessoas por recurso à análise de imagens, sustenta a viabilidade e a aplicabilidade da técnica ICA ao reconhecimento de imagens. Sendo uma técnica estatística de ordem superior, é mais robusta e firme que os métodos de segunda ordem pois tem em consideração a informação dos dados contida nos momentos de ordem superior a dois, permitindo a resolução de casos mais complexos.

O trabalho desenvolvido passou pela concepção de um sistema automático de visão, cuja principal finalidade é controlar de forma automática o acesso a áreas restritas, evitando a entrada de “impostores”.

Num sistema real a dimensionalidade e a complexidade da base de dados é um entrave a uma decisão rápida, por este motivo, neste trabalho investiu-se na construção de uma base de dados com apenas uma vista frontal de cada indivíduo, e o reconhecimento foi efectuado utilizando várias imagens de teste extraídas de uma sequência de vídeo captada durante a aproximação do indivíduo à câmara.

Dado que a captação da sequência vídeo foi efectuada com poucas restrições relativamente a condições de iluminação e de movimentação dos indivíduos, na maioria dos casos os mesmos não se apresentavam em pose frontal perante a câmara. Por este motivo, foi incorporada no trabalho a síntese de vistas frontais.

No que respeita à síntese de imagens, o estudo efectuado e descrito no terceiro capítulo permitiu concluir que nem todas as técnicas produzem uma nova vista do mesmo objecto, recorrendo-se por isso neste trabalho ao método *View-Morphing*.

Os resultados apresentados no capítulo quatro, relativos à comparação das taxas de identificação utilizando as imagens extraídas da sequência vídeo e utilizando as imagens sintetizadas a partir destas, sustentam a decisão de recorrer às imagens sintetizadas para efectuar a autenticação de clientes, em detrimento das imagens que lhes dão origem.

Com vista a melhorar o desempenho da técnica ICA e aumentar taxa de autenticação foi efectuado um pré-processamento das imagens, utilizando uma transformação fotométrica e uma normalização geométrica que permitiram compensar pequenas distorções geométricas e diminuir o efeito das variações da luminosidade.

Na verdade, os resultados obtidos fundamentam que se efectue a normalização geométrica e fotométrica das imagens, pois desta forma alcança-se uma taxa de autenticação de 100%, a qual é reduzida para 68% no caso de não se normalizar as imagens.

No que concerne à rejeição de impostores, os resultados obtidos nos testes do tipo “*Leave-One-Out*” efectuados permitem concluir que é possível impedir a entrada destes em áreas de acesso restrito.

Numa valoração mais qualitativa dos processos utilizados e dos resultados obtidos, pode-se constatar uma certa falta de especificidade na identificação e autenticação de

peçoas, o que reduz a robustez dos experimentos realizados. Esta apreciação baseia-se na observação cuidadosa das distâncias entre representações ICA, as quais são, por vezes, pouco diferentes para a identificação correcta e para outras incorrectas.

A anterior constatação poderá ficar a dever-se, pelo menos em parte apreciável, a várias debilidades nos dados utilizados. Assim, as sequências vídeo são de qualidade muito inferior à que pode ser obtida usando maiores cuidados na selecção do equipamento de aquisição e no condicionamento do ambiente, em particular com respeito a iluminação. Por outro lado, a síntese de vistas frontais recorreu a um número reduzido de pontos correspondentes e, de forma ainda mais gravosa, a sua distribuição na face foi demasiado concentrada no eixo médio vertical, conduzindo a vistas apenas realistas numa fracção reduzida da área total das imagens.

Por outro lado, é importante reconhecer que, apesar das limitações apontadas, os resultados quantitativos atingidos são muito prometedores, em especial se tiver em consideração que a superação de muitas das referidas limitações não acarreta grandes dificuldades.

## 6.2 Perspectivas Futuras

No trabalho apresentado nesta dissertação teve-se por objectivo a simulação de um sistema automático de visão para autenticação. Nesta simulação várias etapas antecedentes ao processo de autenticação foram efectuadas manualmente, sendo por isso deixada em aberto a possibilidade de automatizar o sistema.

Os resultados obtidos e a qualidade das imagens extraídas da sequência vídeo indiciam que o condicionamento do ambiente para a aquisição das imagens da base de dados, bem como as da sequência vídeo, deve ser analisado e ponderado com maior detalhe. Por exemplo, seria aconselhável captar a sequência vídeo com uma câmara que permitisse efectuar focagem e *zoom* progressivos, acompanhando a aproximação dos indivíduos à mesma. Deste modo, as imagens relativas a posições mais afastadas da

câmara seriam captadas com maior resolução, o que acresceria a qualidade das imagens sintetizadas. A aquisição das imagens da base de dados deveria ser efectuada sobre um fundo mais uniforme e de preferência mais escuro do que o considerado neste trabalho, sendo ainda desejável que a sequência vídeo fosse captada com um fundo semelhante.

Após recolhida a sequência de vídeo, e como referido na secção 4.3, foi necessário proceder ao recorte das imagens extraídas de forma a eliminar informação não específica à face. Seria interessante segmentar as faces utilizando um método de detecção das mesmas.

No que concerne à síntese de vistas frontais, uma observação das vistas sintetizadas sugere que a qualidade das mesmas deveria ser melhorada, não só utilizando para o efeito imagens com qualidade superior como também através da definição de mais pontos de correspondência entre as imagens e de uma distribuição mais uniforme destes, por forma a cobrir uma zona da face mais vasta. Deste modo, deixa-se em aberto e como sugestão a implementação de um algoritmo de detecção automática de pontos de correspondência entre imagens faciais que representem vistas diferentes do mesmo indivíduo.

A detecção automática dos pontos de controlo utilizados na normalização geométrica a que as imagens foram sujeitas seria também um factor a explorar e implementar.

# REFERÊNCIAS

- [Abrantes\_03]      Abrantes, S. A. - Apontamentos da Teoria da Informação, *Departamento de Engenharia Electrotécnica e de Computadores, FEUP*, 2003.
- [Bartlett\_01]      Bartlett, M. S.; Movellan, J. R.; Sejnowski, T. J. - Face Recognition by Independent Component Analysis, *IEEE Transactions on Neural Networks*, May 2001.
- [Bartlett\_98]      Bartlett, M. S.; Lades, H. M.; Sejnowski, T. J. - Independent Component Representations for Face Recognition, Proceedings of the SPIE Symposium on Electronic Imaging: Science and Technology, vol. 3299, pages 528-539, San Jose, California, January 1998.
- [Beier\_92]      Beier, T., Neely, S. – Feature-Based Image Metamorphosis, In *Proc. SIGGRAPH 92*, pages 26-31, 1992.

- [Beymer\_93] Beymer, D. J. - Face recognition under varying pose, In *A. I. Memo, MIT AI Lab*, no. 1464, 1993.
- [Beymer\_93a] Beymer, D. J.; Shashua, A.; Poggio, T. - Example based image analysis and synthesis, In *A. I. Memo, Artificial Intelligence Laboratory, MIT*, n.º 1431, 1993.
- [Chen\_91] Chen, S. E.; Williams, L. - View Interpolation for Image Synthesis, In *Proc. SIGGRAPH 93*, pages 279-288, 1993.
- [Comon\_94] Comon, P. - Independent component analysis, A new concept?, *Signal Processing*, vol. 36, 287-314, 1994.
- [Draper\_03] Draper, B. A. [et al.] - Recognizing faces with PCA and ICA, *To appear in Computer Vision and Image Understanding*, 2003.
- [Feng\_00] Feng, G. C.; Yuen, P. C. - Recognition of Head-&-Shoulder Face Image Using Virtual Frontal-View Image, *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Cybernetics*, vol. 30, no. 6, 2000.
- [Hajnal\_01] Hajnal, J. V, Hill, D. L. G, Hawkes, D. J. - Medical Image Registration, *CRC press*, 2001.
- [Hartley\_99] Hartley, R., Zisserman, A. - Multiple View Geometry, *CVPR*, June, 1999.
- [Hyvärinen\_99] Hyvärinen, A. - Survey on Independent Component Analysis, *Neural Computing Surveys*, vol. 2: 94-128, 1999.
- [Hyvärinen\_01] Hyvärinen, A.; Karhunen, J.; Oja, E. - Independent Component Analysis, *John Wiley & Sons, Inc.*, New York, first edition, 2001.

- [Jahne\_95] Jahne, B. - Digital Image Processing (third edition): Concepts, Algorithms, and Scientific Applications, *Springer-Verlag*, 1995.
  
- [Lam\_98] Lam, K.; Yan, H. - An Analytic-to-Holistic Approach for Face Recognition Based on a Single Frontal View, *IEEE Transactions Pattern Analysis Machine Intelligence*, vol. 20, no. 7, pages 673-686, 1998.
  
- [Lanitis\_98] Lanitis, A.; Taylor, C. J.; Cootes, T. F. - Automatic Interpretation and Coding of Face Image Using Flexible Models, *IEEE Trans. Pattern Analysis Machine Intelligence*, vol. 20, no. 7, pages 673-686, 1998.
  
- [Nastar\_97] Nastar, C.; Maghaddam, B.; Pentland, A. - Flexible Images: Matching and Recognition Using Learned Deformations, *Computer Vision and Image Understanding*, vol. 65, no. 2, pages 179–191, 1997.
  
- [Pestana\_03] Pestana, H., Gageiro, J. – Análise de dados para ciências sociais. A complementaridade ao SPSS, *Edições Sílabo*, 2003.
  
- [Seitz\_96] Seitz, S. M.; Dyer, C. R. - View Morphing, In *Proc. SIGGRAPH 96*, pages 21-30, 1996.
  
- [Seitz\_95] Seitz, S. M.; Dyer, C. R. - Physically-valid View Synthesis by Image Interpolation, In *Proc. IEEE Workshop on Representations of Visual Scenes*, pages 18-25, 1995.
  
- [Seitz\_97] Seitz, Steven Maxwell - Image-Based Transformation of Viewpoint and Scene Appearance, *PhD Thesis, University of Wisconsin–Madison*, Madison, 1997.

- [Shannon\_48] Shannon, C. E. - A Mathematical Theory of Communication, *The Bell System Technical Journal*, vol. 27, pages 379-423, 623-656, July, October, 1948.
- [Silva\_04] Silva, Susana Ferreira - Reconhecimento de Imagens por meio da Análise de Componentes de Imagem, *MSc thesis, Faculdade de Ciências e Faculdade de Engenharia da Universidade do Porto*, Porto, Abril 2004.
- [Taub\_71] Taub, H., Schilling, D. L. - Principles of Communication Systems, *McGraw-Hill Book Company*, 1971.
- [Trucco\_98] Trucco, E., Verri, A. - Introductory Techniques for 3-D Computer Vision, *Prentice Hall*, 1998.
- [Ullman\_91] Ullman, S.; Basri, R. - Recognition by Linear Combination of Models, *IEEE Transactions Pattern Analysis Machine Intelligence*, vol. 13, no. 10, pages 992-1006, 1991.
- [Zhang\_02] Zhang, J., Dong, H. - CS 223B Computer Vision: Final Project - Image and View Morphing. March, 2002.  
<http://www.stanford.edu/~kaifan/cs223/main.html>

# APÊNDICE A - NOTAÇÃO

Na tabela apresenta-se, para maior clareza de leitura, a notação utilizada ao longo do texto:

**Tabela A.1 - Notação utilizada e respectiva descrição**

| Descrição                 | Estilo de letra                              |
|---------------------------|----------------------------------------------|
| Escalaes ( $a$ )          | Estilo itálico, letra minúscula              |
| Vectores ( $\mathbf{x}$ ) | Estilo negrito, letra minúscula              |
| Matrizes ( $\mathbf{X}$ ) | Estilo negrito, letra maiúscula              |
| Funções ( $f$ e $F$ )     | Estilo itálico, letra minúscula ou maiúscula |
| Imagens ( $I$ )           | Estilo itálico, letra maiúscula              |