ESPELHOS MATRICIAIS

Renato Roque

2009

FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Mestrado Multimédia

ESPELHOS MATRICIAIS

Reconstruir retratos humanos com componentes estatísticos

Orientação do Professor Dr. Francisco José Lage Campelo Calheiros - Dep. Engª. Civil da FEUP
Co-orientação do Dr. João Manuel de Sousa Duarte Fernandes - Director do Museu de Serralves

Submetida para satisfação parcial dos requisitos do grau de Mestre em Multimédia

Renato Roque
Outubro 2009

Tese de Mestrado Multimédia – FEUP

Título: Espelhos Matriciais

Autor: Renato Roque

Licenciatura em Engenharia Electrotécnica – Telecomunicações – FEUP

Mestrado em Sistemas Digitais e Computadores – FEUP

Orientador: Professor Doutor Francisco José Lage Campelo Calheiros, Dep. Engª Civil FEUP

Co-orientador: Dr. João Manuel de Sousa Duarte Fernandes, Director do Museu de Serralves

Desenhos: *Mente Semente de Cristina Valadas*

Textos: Elliot Rain, *Filipa Leal, Gonçalo M. Tavares, João Fernandes, Jorge Velhote, Luís Tobias, Marta Bernardes, Reborn Bramble, Regina Guimarães, René Ronce*

Outubro 2009

A face humana é igual à daqueles deuses orientais: várias faces sobrepostas em diferentes planos, e é impossível vê-las todas de uma só vez.

Marcel Proust

O teu sorriso permanente

Para o André Puga

Um dia, há muito tempo, um amigo levou-me a visitar um país novo, desconhecido.

Tu devolveste-me esse entusiasmo antigo e quase esquecido de uma nova partida para uma viagem mágica.

Ele deu-me a mão e guiou-me. Tu deste-me a tua mão e guiaste-me.

Ele levou-me a uma clareira no meio de uma floresta. Contigo, entrei num mundo misterioso feito de caixas: caixas dentro de caixas, caixas de vários tamanhos, caixas que pareciam todas iguais.

À primeira vista pareciam todas vazias. "É necessário habituar o olhar para sermos capazes de eliminar as redundâncias que nos inundam a retina e nos impedem de ver o que é importante", aconselhavas-me tu, com paciência.

Eu escutava e perguntava-te "Tal como no Principezinho?" e tentava imaginar quantas formas diferentes existiriam para dizer "ver com o coração".

Tu insistias "Tens de fixar o olhar e ser capaz de isolar as coincidências suspeitas. Se o conseguires, descobres que as caixas não estão vazias, mas habitadas".

Semicerrei as pálpebras e, depois de abrir e fechar muitas caixas, percebi que lá dentro residiam afinal pequenos seres fantasmagóricos, quase invisíveis, como que feitos de bruma.

Eram seres estranhos e mágicos que estavam dispostos a conceder-nos um desejo. Um só desejo.

E tu pediste "Quero sorrir toda a minha vida. Aconteça o que acontecer, quero sorrir sempre", e o fantasma-mor, que habitava na maior caixa de todas, concedeu-te o desejo: um sorriso constante que se revelou à medida que foste fechando as caixas, tal como um retrato numa folha de papel branco impressionada pela luz, depois de mergulhada no revelador.

E eu pedi "Quero ter um amigo". E o fantasma obediente também concedeu o meu desejo.

Mas, o fantasma era um *filho da mãe* e enganou-nos. E, agora, eu não consigo libertar-me do teu sorriso. O teu sorriso permanente escorre de todas as caixas e molha o ecrã do computador.

O teu sorriso revelado e fixado. Aquele teu sorriso malandro de quem tudo está a perceber, mas não quer dizer. Que só diz, se lhe perguntarem. O sorriso de quem parece ser capaz de fintar o mundo.

O fantasma fintou-nos. Agora, o teu sorriso permanece no ar como o sorriso do gato de Cheshire.

> – Gostava que não estivesses sempre a aparecer e a desaparecer assim tão de repente – disse Alice
> – Está bem – disse o gato e, desta vez, desapareceu muito devagarinho, começando pela cauda e acabando com o sorriso, que ainda ficou a ver-se durante algum tempo depois do resto já ter desaparecido."

Por isso, talvez, pior do que perder um amigo é perder alguém que sabemos iria ser nosso amigo. E ficar só com o seu sorriso húmido no regaço. E repararmos que, só hoje, tão tarde, temos a coragem de nos tratar por tu…

Renato Roque, 17 de Junho de 2009

Agradecimentos

Este trabalho é dedicado ao André Puga. Grande parte do que escrevi não teria nunca sido possível sem a sua orientação. Quando em Março de 2008 lhe falei a primeira vez sobre a possibilidade de orientar a minha tese, eu tinha ainda uma ideia muito vaga do que pretendia fazer. Sabia que queria utilizar a fotografia, pretendia concentrar-me no retrato e explorar um conjunto de conceitos da estatística e da teoria da informação, pensando desde essa altura que o trabalho poderia inspirar projectos fotográficos em torno das questões da identidade. O André Puga recebeu-me com uma grande simpatia, com aquele sorriso inteligente que deixou a molhar o meu regaço quando partiu e, com uma enorme tranquilidade, aconselhou-me a não me precipitar e a pensar durante alguns meses sobre o que pretenderia fazer. Deu-me para ler o trabalho de dissertação da Marian Bartlett (1998). Essa tese acabaria por ser uma das minhas leituras de férias, no verão de 2008. Tenho de reconhecer que, tendo na altura apreendido o essencial desse trabalho, me escapavam muitos passos, por me faltar a formação de base na área: PCA e ICA eram para mim domínios completamente novos. Já em Setembro de 2008, regressado às lides na FEUP, voltei a falar-lhe, completamente decidido a tê-lo como orientador, e só então soube da doença, mas o seu optimismo e vontade de fazer coisas levaram-me a acreditar que a doença tinha sido vencida. E André Puga foi, desde que iniciei em Outubro de 2008 este trabalho, um verdadeiro orientador, propondo caminhos, evitando algumas vezes a minha tendência para aprofundar tudo antes de experimentar, ouvindo sempre os meus pontos de vista, atento aos meus objectivos, às minhas dificuldades, para encontrar comigo o caminho certo para prosseguir o trabalho. A questão que posso colocar hoje é se este trabalho fará algum sentido sem ele?

Mas para o sucesso deste trabalho muitas outras pessoas deram contributos fundamentais. As palavras seguintes pretendem expressar de uma forma simples o meu agradecimento a todas elas.

Agradecer ao João Fernandes que aceitou de imediato ajudar-me como co-orientador na área artística e fotográfica.

Agradecer ao professor Fernando Barbosa e ao Pedro Almeida da Faculdade de Psicologia por terem revisto, comentado, feito sugestões para a tese e por terem aceitado ser co-autores de dois artigos escritos durante este trabalho.

Agradecer ao Francisco Calheiros, velho amigo dos bancos da escola, amigo reencontrado agora na FEUP, que decidiu aceitar ser meu companheiro no resto desta viagem pelo reino mágico da fotografia.

Agradecer às 439 pessoas que fotografei: desconhecidos, amigos e familiares que aceitaram participar no projecto e ser retratados na FEUP, no INESC e no ISEP. Sem elas não teria podido construir a BD de retratos que deu suporte a este trabalho.

Agradecer aos colegas e aos professores do Mestrado Multimédia, em especial aos que participaram no trabalho, ao aceitar fazer parte dos 439 retratados e ao responder ao inquérito que utilizei para validar a capacidade de reconhecimento de retratos reconstruídos.

Agradecer ao Eurico Carrapatoso pela paciência com que me aturou e me ajudou, quando foi necessário. Foi mesmo capaz de encontrar tempo para rever com a atenção, que quem o conhece não dispensa, alguns artigos que realizei no âmbito deste trabalho e versões preliminares desta tese.

Agradecer ao Jaime Cardoso pelos comentários e pelas sugestões.

Agradecer ao meu amigo José Manuel Oliveira, pois calhou ser o rosto mais presente neste trabalho. Não foi deliberado; o acaso ditou que a câmara numerasse o seu retrato como 0001 e é por isso o primeiro da BD.

Agradecer à Rosa pela paciência que um trabalho destes exige, mesmo quando pareceu ter esgotado essa paciência.

Agradecer à minha neta Renata por cada segunda-feira em que gaiteira entrou pelo meu escritório dentro, para me obrigar a despegar do monitor, à frente do qual, primeiro decompunha e reconstruía retratos e agora termino esta tese.

Agradecer ao Elliot Rain, ao Emílio Remelhe, ao Jorge Velhote, ao Luís Tobias, à Marta Bernardes, ao Reborn

Bramble, à Regina Guimarães e ao René Ronce por terem respondido ao meu desafio e terem escrito textos inéditos para esta tese.

Agradecer à Filipa Leal e ao Gonçalo M. Tavares por me terem oferecido para este trabalho textos seus, ainda que não inéditos.

Agradecer à Cristina Valadas pelos belíssimos desenhos. Quando vi os desenhos e o texto do Emílio Remelhe no catálogo da exposição ***Mente Semente***, imediatamente os associei ao que andava a fazer. Queria utilizá-los para ilustrar este trabalho. Agradeço-lhe ter tornado esse desejo possível.

Agradecer ao Rui Mendonça pelas fotografias dos desenhos da Cristina e por muitas palavras amigas que me ajudaram na formatação deste trabalho.

Agradecer a muitos amigos a quem fui enviando versões preliminares da tese e que me foram enviando comentários: o António, o Augusto, o Renato, a Rita.

Renato Roque, Julho 2009

ÍNDICE

1 INTRODUÇÃO 9

1.1 Porquê espelhos 9

1.2 Objectivos do trabalho 11

1.3 Motivações para o trabalho 12

1.4 Estrutura deste documento 12

2 PILARES ESSENCIAIS DO NOSSO TRABALHO 17

2.1 O retrato e a identidade na fotografia 17

2.2 As teorias da percepção 25

2.3 Modelos de computação e processamento de imagem 37

3 METODOLOGIA 61

3.1 Construção da BD de retratos 61

3.2 Conceito de erro de reconstrução 62

4 RECONSTRUÇÃO DE RETRATOS COM A BD INICIAL 67

4.1 A BD de retratos 67

4.2 Resultados com o PCA 68

4.3 Resultados com o ICA 80

4.4 Resultados com o PCA + ICA 96

4.5 Resultados com o NMF 99

4.6 Resultados com PCA + ICA + NMF 120

4.7 Conclusões sobre as diferentes técnicas estatísticas 122

5 RECONSTRUÇÃO DE RETRATOS COM A BD NORMALIZADA 143

5.1 Comparação entre as duas BDs para todos os retratos da BD 145

5.2 Comparação entre as duas BDs para todos os retratos fora da BD 151

5.3 Erro fotográfico para as duas BDs 153

5.4 Conclusões sobre o uso da BD normalizada 155

6 ANÁLISE DOS RESULTADOS DOS INQUÉRITOS 159

6.1 Organização do inquérito 159

6.2 Resultados do inquérito 161

6.3 Avaliação dos resultados do inquérito com a SDT 162

7 ALGORITMOS PARA A INTEGRAÇÃO DUAS BDS DE RETRATOS 169

7.1 Modelo de memória 169

7.2 Algoritmo inicial 170

7.3 Algoritmo com número global de componentes imposto 172

7.4 Conclusões 177

8 PROJECTO ARTÍSTICO 181

9 CONCLUSÕES 185

9.1 Conclusões gerais 185

9.2 Conclusões para o NMF 187

9.3 Conclusões sobre algoritmos de integração de duas BDs 187

10 ANEXOS 191

10.1 Endereços URL 191

10.2 Referências 192

10.3 Conceitos da teoria da informação 194

10.4 Inquérito 196

10.5 Texto para projecto fotográfico 205

10.6 Publicações realizadas 207

1 Introdução

Este trabalho tem como ponto de partida a fotografia, em particular o retrato.

Aborda um conjunto de questões relacionadas com a percepção visual humana, em particular a percepção de imagens de faces.

Utiliza técnicas de computação e de processamento de imagem para decompor retratos de faces humanas em componentes principais, componentes independentes e componentes positivos, para avaliar e comparar essas técnicas na sua capacidade de reconstrução dos retratos, em particular com muito poucos componentes.

1.1 Porquê espelhos

O processo fotográfico tem cerca de 150 anos de idade. Quando surgiu foi encarado sobretudo como um processo rigoroso para registo da realidade. Daguerre, um dos pioneiros, no século XIX, nos primeiros tempos da fotografia, quando pretendia explicar essa nova técnica extraordinária, descrevia-a como uma espécie de **espelho com memória**, ou seja como um processo capaz de reflectir o real mas também capaz de o capturar, registando-o de uma forma permanente. Não é por acaso que o 1º livro de fotografia de H.Fox Talbot (1844) se chama "The Pencil of Nature".

Figura 1 – The Pencil of Nature - o 1º livro de fotografia

Essa capacidade mágica para parar o tempo, para congelar uma fracção do segundo, como um espelho com memória, continua a ser talvez a característica mais diferenciadora da fotografia, mas também porventura uma das suas maiores fontes de equívocos.

Curiosamente também, a palavra espelho deriva do latim ***speculum***, que tem a mesma origem da palavra portuguesa especular, que na antiguidade significava observar os céus e os movimentos das estrelas com recurso a espelhos, e que hoje se utiliza para meditar, procurar a solução para um problema através do uso da razão. Paradoxalmente talvez, ou não, a fotografia contemporânea parece hoje reencontrar muitas vezes estes dois conceitos: reflectir o mundo (espelho) e utilizar a razão na procura de alguma coisa (especular).

Continuando nesta nossa procura de razões – como se razões fossem necessárias para termos escolhido *Espelhos Matriciais* como título da tese – poderíamos realçar ainda que o verbo latino *spectare*[1] significa ver, olhar, apreciar e que a palavra portuguesa espectro deriva de *spectrum*, que em latim curiosamente significa imagem, visão. Ora a decomposição espectral da matriz de covariâncias, construída a partir dos retratos, é um dos suportes essenciais da técnica de Principal Component Analysis (PCA), a primeira técnica estatística que iremos experimentar e validar. E o PCA é, de alguma forma, um suporte de base para as demais técnicas estatísticas. Ora todas essas técnicas conduzem, quando utilizadas em retratos, como veremos, à descoberta de um conjunto de retratos espectrais, chamados componentes. Descobrimos assim matrizes 'mágicas' que 'escondem' dentro de si todas as características dos retratos originais. Descobrimos assim espelhos matriciais, espelhos misteriosos que parecem deformar as imagens originais, tal como os espelhos não-planos da óptica, mas onde a deformação é apenas uma ilusão, pois permitem a reconstrução perfeita dos retratos. Como se, ao utilizarmos essas técnicas estatísticas encontrássemos de alguma forma as imagens, os retratos (*spectra*) essenciais, que nos permitem reconstruir todos os retratos: a pedra filosofal do retrato humano

A somar a estas razões "**espelho**" tem sido uma espécie de palavra mágica para o autor, palavra que, de uma forma explícita ou implícita, tem aparecido em quase todos os seus projectos fotográficos, e que simboliza a fotografia e a procura pessoal que a fotografia propicia.

Toda a fotografia é um auto-retrato
(Boris Lehman 2003)

Figura 2 – Auto-retrato no espelho

Poderíamos enumerar os projectos "*Entre dois espelhos a 10º*", "*10º para a meia-noite*" com os espelhos do projecto anterior a fecharem-se, "*Espelhos*", "*Novos Espelhos*", etc.

Como razão última poderíamos invocar que a tese, e em particular o projecto artístico associado, se vai debruçar sobre questões de **identidade**, identidade quase sempre simbolizada no retrato por uma imagem da face, consubstanciada na fotografia do BI ou do passaporte, imagem que afinal o homem conhece desde há muito tempo do espelho, como única forma que antes tinha de se olhar e de se conhecer ao longo da vida.

No fundo vamos lidar neste projecto com **espelhos especiais** que reflectem e memorizam, mas ao mesmo tempo eliminam redundâncias e permitem descobrir e guardar o essencial. O essencial? Esse pretende ser afinal um ponto de reflexão.

(Barthes 1980); (Roque 1986); (Roque 1988); (Roque 1992, 1992); (Frizot 1998); (Valadas 2008);

1 Em rigor dever-se-ia dizer o verbo *specto, spectas, spectavi, spectatum*

1.2 Objectivos do trabalho

O trabalho realizado é construído num 'espaço 3D', definido por 3 vectores quase ortogonais:

- A fotografia
- A computação e processamento de imagem
- A psicologia

Nesse espaço 3D, o ponto central para o nosso trabalho situa-se mais na proximidade do eixo da fotografia e do eixo da computação e processamento de imagem. A fotografia é um eixo dominante do nosso trabalho. O trabalho vai ser suportado numa BD de retratos frontais, construída por nós.

A computação e processamento de imagem, como segundo eixo do nosso trabalho, vai disponibilizar um conjunto de ferramentas estatísticas, que pretendemos avaliar e comparar. Ao contrário de outros trabalhos de investigação na área do reconhecimento de faces, que têm utilizado essas técnicas estatísticas para tentar aumentar a rapidez ou a eficiência de algoritmos, ou propor novos algoritmos, em particular para o reconhecimento automático de imagens de faces, a nós interessa-nos sobretudo a capacidade desses processos, não para reconhecer automaticamente a identidade de uma pessoa, mas para reconstruir retratos fotográficos reconhecíveis por humanos. Vamos utilizar a base de dados de retratos que construímos para comparar os erros geométricos e fotográficos para as diferentes técnicas estatísticas na reconstrução de um retrato e para avaliar a forma como esse erro evolui com o número de componentes utilizados. Para além de avaliar e comparar diferentes técnicas estatísticas, pretendemos também construir e avaliar sistemas híbridos, que associam técnicas diversas, por exemplo PCA+ICA e PCA+ICA+NMF, na reconstrução fotográfica. Pretendemos também avaliar algoritmos que permitam a integração de novos retratos numa BD de conhecimento de retratos, assentes nas técnicas estatísticas avaliadas, fazendo uma integração optimizada de componentes, permitindo dessa forma uma "aprendizagem" de novos retratos.

O eixo da psicologia interessa-nos na medida em que constatámos que quase todo o trabalho nestas áreas tem sido suportado por um trabalho teórico e experimental na área da psicologia e da neurologia. Nesta fronteira pretendemos em particular contribuir para de alguma forma validar a ideia de que pode haver nos processos de percepção visual mecanismos que permitem construir imagens reconhecíveis, em particular de rostos humanos, com informação esparsa, ou seja com um pequeno número de componentes, optimizando assim o processo, nomeadamente do ponto de vista da quantidade de informação e da energia necessária. Essa validação vai ser completada no plano fotográfico, com inquéritos, utilizando retratos reconstruídos por nós com poucos componentes e com valores diversos para o erro, procurando assim de alguma forma simular e validar processos esparsos de reconstrução e reconhecimento de retratos.

Os objectivos definidos atrás irão ser consolidados através de um trabalho que se desenvolve num outro 'espaço 3D', segundo 3 componentes:

- **Desenvolvimento** – em particular de um conjunto de scripts em Matlab para normalização da BD, para tratamento estatístico dos retratos da nossa BD, utilizando as diferentes técnicas que experimentámos no nosso trabalho e para desenvolver um conjunto de algoritmos de suporte ao nosso trabalho.
- **Investigação** – validação de soluções esparsas de construção de retratos, com diferentes técnicas estatísticas e com diferentes erros, para o reconhecimento humano de retratos e validação de um modelo para integração de dois conjuntos de imagens, mediante uma integração optimizada de componentes estatísticos do mesmo tipo ou de tipos diferentes.
- **realização de um projecto artístico** - os resultados deste trabalho servirão de base para desenvolver um projecto fotográfico, que pretende ser uma reflexão pessoal em torno do retrato humano e das questões que os métodos estatísticos estudados poderão colocar, na área do retrato e da identidade e de como nós a percepcionamos.

1.3 Motivações para o trabalho

Quase sempre, ao longo da história da fotografia desde a sua invenção no século XIX, que ela tem sido encarada como um processo que envolve mistério e magia. Mesmo hoje, com a banalização da imagem fotográfica, essa aura de magia da fotografia está longe de ter desaparecido.

Ser confrontado com técnicas estatísticas, capazes de decompor retratos em componentes fantasmagóricos ou abstractos, que depois permitem, se somados nas proporções correctas, reconstruir de uma forma perfeita todos os retratos de uma BD e mesmo, com um pequeno erro, novos retratos, constituiu um novo mistério a somar ao mistério da fotografia, um mistério que ainda hoje, depois de muitos meses a trabalhar com todas estas técnicas estatísticas, nos surpreende e maravilha.

Encher uma tina de componentes PCA ou ICA e nela mergulhar uma folha de papel em branco, onde ao fim de algum tempo, por acção dos componentes, um retrato se revela, foi o regresso mágico à Câmara Escura que descobrira há tantos anos.

> *"Ia visitar um país novo, desconhecido. Na hora da partida estava entusiasmado, como sempre estou, ao partir de viagem.*
>
> *O Luís deu-me a mão e guiou-me. Andámos durante muito tempo. Atravessámos campos e aldeias, prados e bosques. Perguntei-lhe várias vezes se ainda demorava muito. Ele respondia-me cheio de indulgência, que eu precisava de ter paciência.*
>
> *Entrámos por fim numa clareira no meio de uma floresta. Parámos. Algo na expressão do Luís me dizia que tínhamos chegado. Na clareira havia uma pequena casa. À primeira vista, esta casa nada possuía de extraordinário, mas à medida que os meus olhos se adaptaram à luminosidade da clareira, notei que era rodeada por uma espécie de brilho, talvez causado pela luz solar filtrada pela folhagem das árvores que a cobriam. Aproximei-me, andei à sua volta. A casa estava fechada. As janelas eram cobertas por portadas grossas. Nada se via para o interior. Olhei para o Luís como que a perguntar o que fazer. Ele sorriu, aquele sorriso confiante que só ele sabe fazer, e mandou-me espreitar pelo buraco da fechadura. Caminhei para a porta. Inclinei-me, algo receoso, e espreitei. A princípio nada vi. Os meus olhos, feridos pela claridade do exterior, nada conseguiam ver. Mas a pouco e pouco os contornos foram-se definindo. Um mundo maravilhoso foi-se desenhando através daquele pequeno orifício. Um mundo, onde as regras a que nos habituamos e que acabamos por respeitar podem desaparecer. Um mundo de pernas para o ar! Fiquei fascinado. O meu olho direito cansado, colado ao buraco da fechadura daquele mundo (i)rreal começou-me a doer. Mas poucas vezes há maravilha sem um pouco de dor e de sacrifício. O Luís chamou-me. Tínhamos de voltar. Eu não queria partir. O Luís explicou-me então que poderia levar aquela casa comigo. Podia levá-la dentro do coração. O Luís falou-me como só ele consegue falar coma sua voz pausada, segura, amiga. Depois de o ter ouvido, sabia que era verdade. Ergui-me e olhei para ele. Ele sorriu de novo e percebi que a casa já estava dentro de mim. O seu sorriso era maravilhoso..."*
>
> *Em **Elogio da Fotografia**, Renato Roque*

E se pensarmos que talvez haja no nosso cérebro mecanismos equivalentes, que nos permitem optimizar o processo de reconhecimento de rostos humanos, mais maravilhados teremos de ficar.

Por um lado a experiência na fotografia desde os anos 80 e por outro lado a formação técnica em engenharia criaram porventura as condições para realizar um trabalho, se calhar pouco comum, num território difícil e inóspito, por estar na fronteira entre a arte e a fotografia e as técnicas estatísticas de computação e de processamento de imagem, para imagens de rostos: um trabalho na fronteira entre a tecnologia/ciência e a arte.

1.4 Estrutura deste documento

Na introdução fez-se uma primeira apresentação deste trabalho, justificando o título escolhido, definindo os grandes objectivos e apresentando as principais motivações. Aqui apresentamos a

estrutura das secções seguintes.

Na secção 2 é feita uma descrição daquilo que habitualmente se designa por estado da arte, apresentando as técnicas em que o trabalho se vai basear e procurando situar qual o ponto em que as áreas que suportam o trabalho se encontram.

Na secção 3 faz-se uma apresentação sucinta da metodologia utilizada na obtenção dos resultados apresentados nas secções 4, 5 e 6.

Na secção 4 apresentam-se e comparam-se os resultados obtidos com a nossa BD de retratos, utilizando as diferentes técnicas estatísticas.

Na secção 5 apresentam-se e comparam-se os resultados obtidos com a nossa BD de retratos depois de normalizada, garantindo o mesmo posicionamento para os olhos em todos os retratos. Os resultados são comparados com os da BD não normalizada da secção 4.

Na secção 6 são apresentados os resultados do inquérito realizado, para avaliar a capacidade de reconhecimento humano de retratos reconstruídos com erro, utilizando um número pequeno de componentes.

Na secção 7 são apresentados e avaliados algoritmos para integrar imagens, mediante a integração optimizada de componentes PCA ou ICA, inspirados num modelo de memória suportado em duas memórias, de curto e de longo prazo, para a aprendizagem não supervisionada de imagens de rostos.

Na secção 8 é apresentado e descrito o projecto artístico a realizar.

Na secção 9 são apresentadas as conclusões

São ainda apresentados 6 anexos na secção 10: o primeiro com os URLs do SW e das BDs utilizadas, o segundo com as referências bibliográficas que suportam o trabalho, o terceiro onde se desenvolvem alguns conceitos da teoria da informação referidos no trabalho, o quarto com um exemplar do inquérito realizado, o quinto com o texto escrito por João Fernandes para o projecto fotográfico a realizar no âmbito desta tese e por fim o último anexo com a lista dos artigos publicados até à data das provas de dissertação desta tese.

Nota sobre referências: procurámos fundamentar este trabalho com muitos trabalhos relevantes, que referimos e identificamos, em todas as áreas que este trabalho aborda. Em vez de fazer citações desses trabalhos, optámos por interpretá-los e sintetizá-los, tendo em conta os aspectos relevantes de cada um para esta tese. Por isso, decidimos colocar a lista das referências importantes para cada secção deste documento no fim dessa secção, em vez de as disseminar ao longo do texto.

Sonho com um espelho. Vejo-me a usar uma máscara ou vejo no espelho alguém que sou eu mas que não reconheço como eu.

Jorge Luís Borges

o retrato é um mal que não tem cura:
uma vez tirado, dura, dura, dura.
diz-se que os primitivos
que cuidavam dos seres vivos
com particular ternura
- isto é: acautelando a sua fragilidade -
não queriam ser retratados
porque se sentiam roubados.
também se diz que um bom retrato
por vezes nos olha
mais do que nós o olhamos
e isso pode ser verdade
a ponto de nos sentirmos vigiados.

a única coisa que sei ao certo
é que o único retrato oficial
que me tiraram na vida
foi fotografado no dia
em que o meu pai morreu.
ele não podia saber
e eu também não.
mas quando olho para esse retrato
vejo que um pedaço de mim partiu
sem eu acenar esse grande lenço
que é toalha de Verónica
e vaso de todas as lágrimas.

Regina Guimarães

2 Pilares essenciais do nosso trabalho

Nesta secção estabelecemos as linhas gerais do que se convencionou chamar de "Estado da Arte", para os vários eixos em que este projecto se desenvolve, erguendo assim os pilares essenciais em que o nosso trabalho irá assentar.

Tentamos começar por fazer uma pequena contextualização histórica das principais referências que nos deram suporte: na fotografia, nas teorias da percepção visual e na computação e processamento da imagem.

- A fotografia – O retrato e a identidade na fotografia – a importância do retrato na história da fotografia que, de alguma forma, assumiu a função antes desempenhada pela pintura no que diz respeito ao retrato. Alguns projectos fotográficos artísticos relevantes em torno do conceito de identidade.
- A psicologia – As teorias da percepção visual – teorias que influenciaram de uma forma decisiva as abordagens mais recentes dos modelos de computação e de processamento de imagem para o reconhecimento de padrões, em particular do rosto humano e para processos de aprendizagem não supervisionada.
- A computação e processamento de imagem – Modelos de computação e de processamento de imagem – modelos propostos para formalizar a representação, o reconhecimento e a aprendizagem de padrões, em particular do rosto humano.nos modelos de computação e de processamento de imagem de rostos

Depois desenvolvemos um pouco mais as principais contribuições experimentais e teóricas para o nosso trabalho.

2.1 O retrato e a identidade na fotografia

Se quiséssemos aqui escrever, nem que fosse uma pequena história resumida da fotografia e em particular do retrato, teríamos com certeza de começar pelos primeiros retratos dos daguerreotipistas e depois por Nadar, se calhar pelo seu célebre retrato da mãe, continuando com um sem número de fotógrafos europeus e americanos, isto se nos quiséssemos limitar à fotografia no Ocidente, até chegarmos porventura ao retrato (des)conhecido da mãe de Barthes, de que ele fala sempre no seu iconográfico livro "Câmara Clara", mas que nunca mostra. Se pensarmos nos fotógrafos significativos do século XX, quase todos fizeram retrato. E depois, se pretendêssemos abordar a chamada fotografia contemporânea, seríamos obrigados a mencionar dezenas ou mesmo centenas de nomes e centenas de projectos fotográficos, todos eles construídos a partir do retrato e colocando questões de alguma forma relacionadas com a identidade. Mas para tal necessitaríamos de tempo e de espaço que este trabalho não comporta. Vamos por isso limitarmo-nos a tentar contextualizar o nosso trabalho e para tal referir um conjunto de projectos e de fotógrafos, escolha que obedece apenas a um critério pessoal, decorrente do gosto ou da pertinência intuída para o nosso trabalho. Recomendamos a quem pretender aprofundar esta história dois livros iconográficos: "Faces" de Ben Maddow e "Face-The new photographic portrait" de William Ewing. O primeiro conta toda a história do retrato, desde a invenção da fotografia até à década de 70 do século XX. O segundo apresenta um conjunto de projectos de retrato contemporâneos considerados mais importantes pelo autor.

É indiscutível a importância do retrato na história da fotografia. Desde que o processo fotográfico foi descoberto em meados do século XIX que a fotografia assumiu um papel fundamental no registo documental. E o retrato tem tido indiscutivelmente um papel de relevo nesse registo. E tal

como acontecera a muitos pintores na história da pintura também muitos fotógrafos recorreram frequentemente ao auto-retrato como uma forma de conseguirem registar-se a si próprios e de assim marcarem o seu trabalho com a sua presença.

Essa capacidade surpreendente que a tecnologia passou a oferecer, de criar um espelho mágico com memória, de registar o que acontecia, de roubar instantes de tempo ao tempo, foi sentida imediatamente pelos pintores como uma verdadeira ameaça e por isso reagiram, muitas vezes violentamente à divulgação e à popularidade crescente da fotografia. Os pintores temiam que a fotografia, pela sua simplicidade e pela sua "veracidade", acabasse com o negócio do retrato pintado a óleo, o que em parte de facto aconteceu.

A evolução da tecnologia e a oferta de equipamentos cada vez mais simples e mais baratos acabou por tornar a fotografia, durante o século XX, num processo de registo popular e generalizado. Quase todas as famílias, até de um nível económico não muito elevado, usaram a fotografia para registar todos os momentos importantes para a vida familiar. Casamentos, baptizados, aniversários e viagens passaram a ser registados com regularidade. Possuo, tal como as minhas irmãs e quase toda a gente da minha geração, um álbum fotográfico, construído com dedicação pelos meus pais, onde posso recordar o meu nascimento, os meus aniversários, as férias longínquas em Figueira de Castelo Rodrigo, os meus avós e tanta outra gente que entretanto desapareceu. Os meus pais tinham mesmo o cuidado, muito popular nesse tempo, de me fazer fotografar, tal como às minhas irmãs, todos os anos, num estúdio de um fotógrafo profissional. Lembro-me dessas sessões na casa *Gualtieri*, na R. de S. António, e do cuidado posto nas roupas e nos penteados.

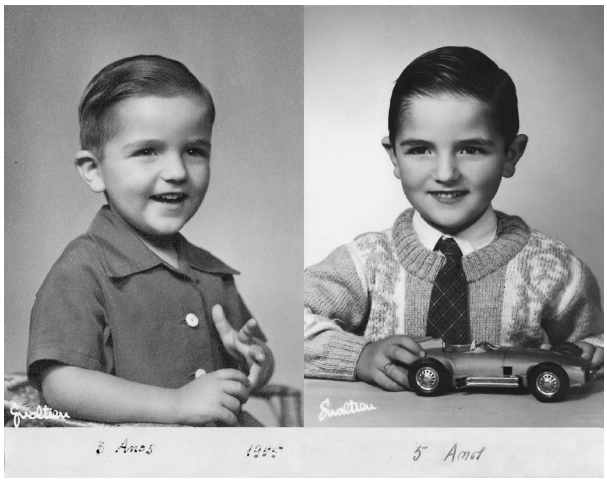


Figura 3 – Retrato aos 3 e 5 anos – casa Gualtieri

A razão de ser da enorme popularidade da fotografia no retrato, destronando claramente a pintura, está com toda a certeza associada à "essência" dessa fotografia, que já era com certeza intuída pelas pessoas, ainda que só dissecada e teorizada nos anos 60, por Rolland Barthes, no seu famoso livro "Câmara Clara". Segundo Barthes a fotografia, muito mais do que a pintura, tem essa capacidade de amarrar a imagem ao referente, ou seja, à pessoa fotografada. Como diz Barthes nesse livro mítico:

> *Na Fotografia a presença da coisa (num determinado momento passado) nunca é metafórica.*
>
> *O passado é a partir de agora tão seguro quanto o presente, aquilo que se vê no papel é tão real como aquilo que se toca.*
>
> *A Fotografia não rememora o passado, confirma que aquilo que vejo existiu realmente*

Por isso, uma percentagem muito significativa, dos muitos biliões de fotografias feitas, desde a sua descoberta no século XIX, foram com toda a certeza retratos.

O retrato também desempenhou um papel central em todos os processos identitários. Nos BIs, nos passaportes, nos cartões de membro de uma qualquer associação mas também nos ficheiros da polícia foi a fotografia de rosto que permitiu identificar cada indivíduo. Há por isso uma ligação muito forte entre o retrato de face e a identidade individual.
Perante a importância documental da fotografia, perante a sua relação com o registo e com a memória, não surpreende a importância do retrato e do auto-retrato na história da fotografia, o que surpreende porventura é a capacidade deste meio "automático e mecânico", como muitos gosta(va)m de caracterizar a fotografia, para se desdobrar em tantas e tão diferentes abordagens do retrato. Desde os retratos poéticos de Cartier-Bresson, que são mais retratos de momentos do que de pessoas, onde as pessoas são meros personagens num fotograma de uma história interrompida, momentos decisivos, como ele gostava de dizer, fragmentos de vida congelados no espaço-tempo, aos retratos surrealistas de Man Ray, às imagens de cores berrantes, de Ouka Lele, que tão bem retratam a Espanha liberta dos anos oitenta, ou até as imagens maneiristas de Mapplehorpe, com conteúdos eróticos e sexuais tão flagrantes.

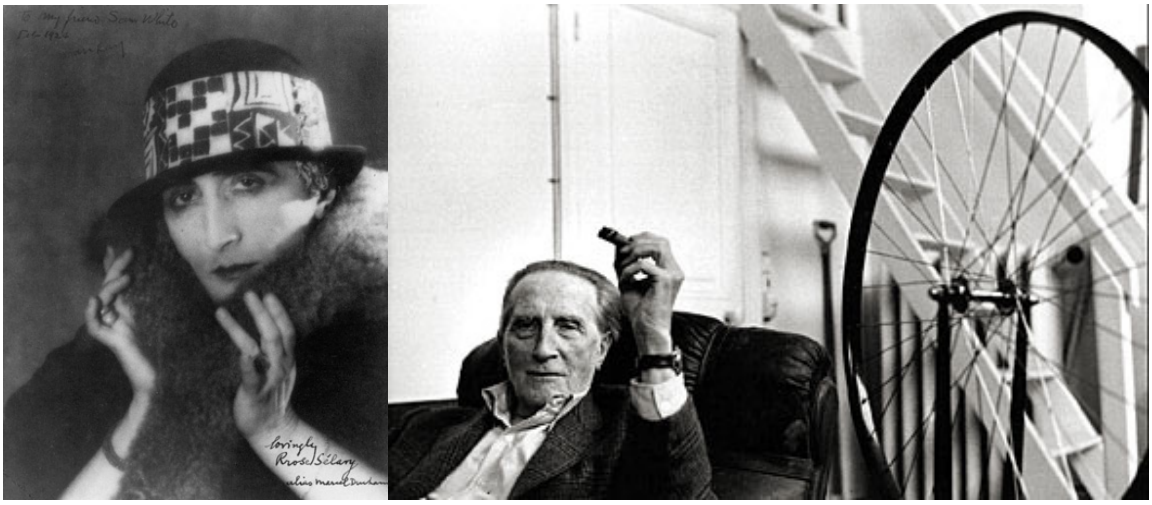


Figura 4 – Rrose Selavy por Man Ray e Duchamp por Henri Cartier Bresson

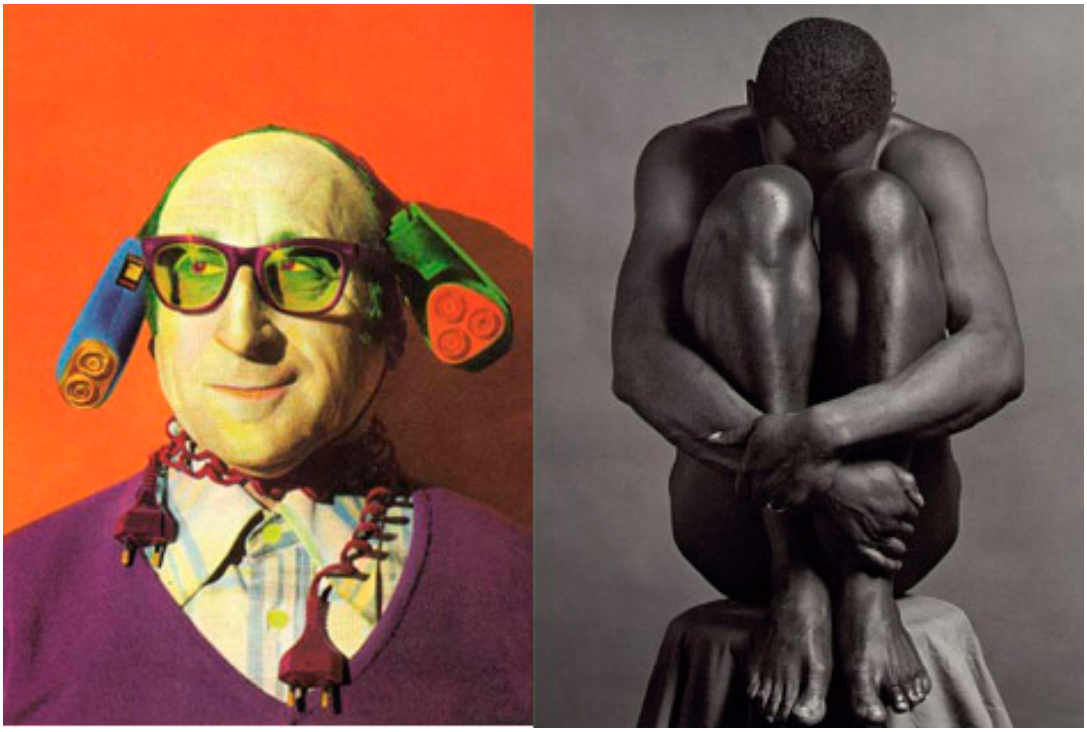


Figura 5 – Retratos de Ouka Lele e Mapplehorpe

Um outro aspecto que nos parece importante realçar é o papel que a fotografia tem desempenhado em muitos projectos de construção de uma espécie de arquivo universal. O projecto **Museu Imaginário** de André Malraux poderá funcionar como paradigma desta visão, talvez megalómana, de conseguir coleccionar e arquivar tudo o que nos rodeia. A fotografia era o meio ideal para tal empreendimento.

Figura 6 – O Museu Imaginário de André Malraux

O museu imaginário de Malraux construía-se juntando fotografias de todos os objectos de arte, criando um museu universal, onde todo o património artístico estaria disponível. Muitos projectos fotográficos e também alguns dos que destacamos nesta curta apresentação têm esta componente arquivística como vector, esta visão de colecção infinita, de recolha e de conservação do mundo através da fotografia.

O retrato tem servido em muitos outros casos para essa tentativa de registo sistemático, quase obsessivo, procurando desta forma uma garantia de memória colectiva. O fotógrafo alemão August Sander é um exemplo paradigmático da missão arquivística associada à fotografia. Sander procurou registar meticulosamente em chapa todos os personagens da sua época, para se tornarem, com o passar do tempo, personagens simbólicos desse tempo.

Figura 7 – Fotografias de August Sander

Se a fotografia até à década de 60 se caracteriza sobretudo pelo seu lado documental, e também no caso do retrato ou do auto-retrato, a partir do conceptualismo fotográfico, que explode nessa década, muitos fotógrafos e muitos artistas plásticos que recorrem à fotografia, rompem de alguma forma com esses padrões, onde o registo "do que foi" era determinante e utilizam a fotografia para colocar em cima da mesa uma série de questões pertinentes na arte contemporânea e entre elas a questão da identidade.

Inúmeros são os projectos contemporâneos relevantes que colocam de muitas formas a questão da identidade. Comecemos por Joan Fontcuberta, fotógrafo catalão que centra quase toda a sua obra em torno das questões da verdade e da mentira, onde as questões da identidade assumem papel relevante. É caso de um dos seus projectos mais paradigmáticos "Fauna", onde inventa um personagem misteriosa, o professor Ameisenhaufen, recorrendo à fotografia, como registo de autenticação da verdade para tornar esse personagem real.

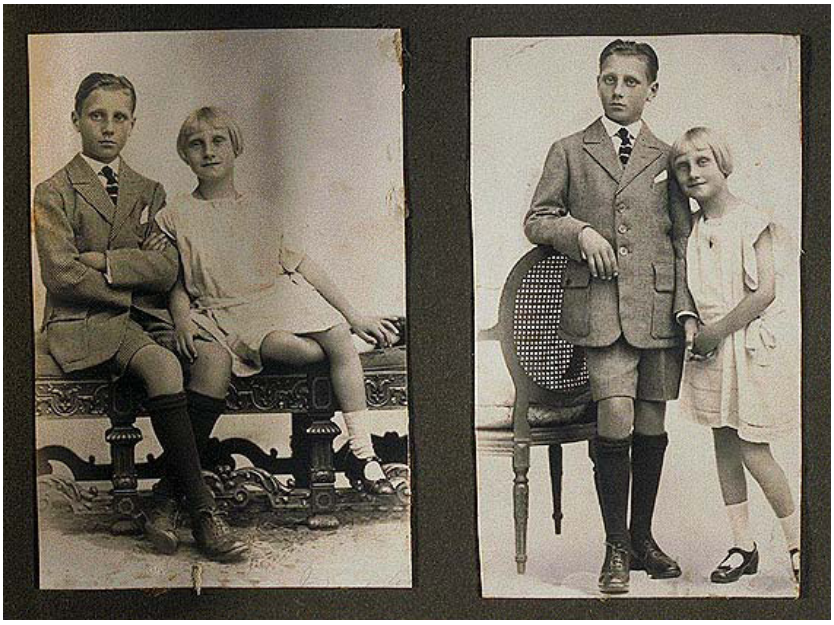


Figura 8 – Retrato de Ameisenhaufen com a irmã, quando adolescentes

Podemos também referir Cindy Sherman, Duane Michaels, Lee Fridlander e Virxilio Viéitez por razões diferentes. Cindy Sherman pela sua notoriedade mediática e por recorrer sempre a auto-retratos encenados onde questiona a sua própria identidade; Duane Michaels pela sua originalidade e pelo seu questionamento permanente de tantas questões éticas, morais e religiosas, nalguns casos a partir da sua homossexualidade assumida; Lee Friedlander, que pertence a uma nova geração de fotógrafos americanos, por ser o autor de um livro iconográfico chamado "Self Portrait", que é uma colecção de auto-retratos; Virxilio Viéitez, que conheci pessoalmente em Braga, um fotógrafo galego que faleceu recentemente, por ter sido durante dezenas de anos "apenas" o fotógrafo do povo das aldeias da região da Galiza onde residia, até ser "descoberto" pela fotografia contemporânea e se tornar num dos fotógrafos espanhóis de maior sucesso, com vendas significativas em galerias de todo o mundo, em especial de Paris. Em virtude da sua idade avançada, era a sua filha que imprimia os trabalhos, encomendados de todo o mundo.

Figura 9 – Cindy Sherman – Still Cuts e retrato de Virxilio Viéitez

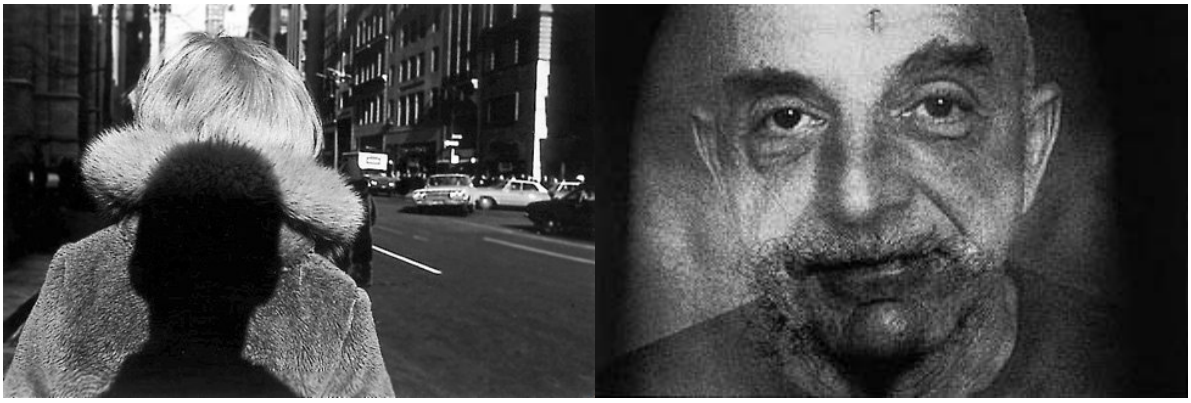


Figura 10 – Auto-retrato de Lee Friedlander e de Duane Michaels

Gostaria ainda de referir Christian Boltansky que tem realizado inúmeros projectos em torno da questão da identidade. Talvez o mais conhecido seja o projecto "Les Suisses Morts", em que utiliza pequenos retratos tipo passe de cidadãos suíços, mortos durante a 2ª guerra mundial, montados numa enorme parede de caixas metálicas, que imediatamente nos evocam um cemitério.

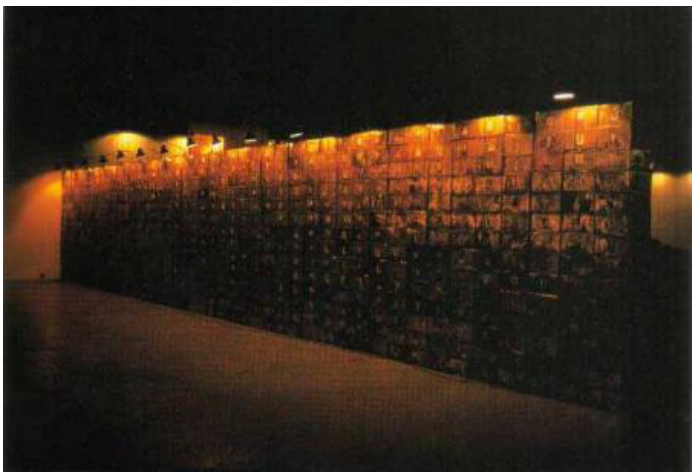


Figura 11 – "Les Suisses morts" de Christian Boltansky

Utilizando a fotografia, Boltanski questiona o seu significado como memória. "Les Suisses Morts" refere-se ao Holocausto, conseguindo ter um enorme impacto emocional, talvez em parte devido aos frágeis materiais que utiliza e que sugerem a fragilidade da vida e as vidas perdidas por aqueles que desapareceram durante o genocídio. Este impacto existe, apesar de sabermos que são imagens de cidadãos suíços e que a Suíça se não envolveu na guerra. Através do uso da fotografia nesta e noutras obras, Boltanski sugere o poder da fotografia em produzir em nós saudade, nostalgia e melancolia pelas vidas dos fotografados, neste caso quer saibamos ou possamos saber que eles pereceram no holocausto, ou não.

Teremos também de referir Thomas Ruff que retrata de forma anódina pessoas anónimas, registadas em fotografias tipo passe, imagens tão idênticas e tão neutras, que parecem transformar todos esses indivíduos num único personagem cinzento e globalizado dos nossos dias. Também se poderia identificar em Ruff uma intenção arquivística, mas no seu caso nesta procura do que de mais neutro existe em todos nós. Ruff também propôs um conjunto de projectos a partir de fotografia tipo passe em que joga com o conceito de identidade, ao sobrepor mais do que uma imagem para conseguir uma nova identidade virtual ou ao jogar com conceitos como género ou idade.

Figura 12 – "Anderes Porträts 1992" de Thomas Ruff

Neste contexto é também quase obrigatória a referência a Gerhard Richter, pela originalidade da sua abordagem. Richter tem utilizado muitas vezes a fotografia como ponto de partida para o seu trabalho de pintura, em particular de retratos, resolvendo assim de outra forma a contradição

fotografia-pintura, colocada desde a invenção do processo fotográfico.

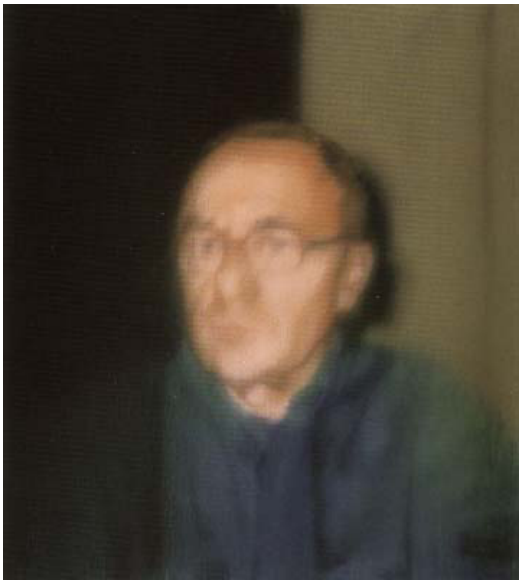


Figura 13 – Auto-retrato de Gerhard Richter

Se olharmos para a fotografia portuguesa contemporânea constatamos facilmente que também muitos dos fotógrafos mais relevantes realizaram muitos projectos, onde a questão da identidade assume papel determinante.

Podemos começar com o Jorge Molder que parece fotografar obcecado em torno do seu próprio retrato e que se auto-fotografa há décadas, das mais variadas formas, sempre a preto e branco.

Figura 14 – Auto-retrato de Jorge Molder e auto-retrato de Helena Almeida

Helena Almeida que começou como pintora e que desde há muito tempo recorre à fotografia para realizar projectos em que ela e o seu corpo são o principal personagem, misturando muitas vezes a fotografia, a pintura e o desenho.

José Luís Neto que realizou um projecto chamado "22474", a partir de uma imagem de Joshua Benoliel, do arquivo fotográfico da Câmara Municipal de Lisboa, realizada em 5 de Fevereiro de 1913. Essa imagem mostra-nos a cerimónia de abolição do capuz na penitenciária de Lisboa. A imagem original foi obtida no anfiteatro desse estabelecimento prisional e consegue-se identificar a cabeça dos presidiários, dentro de cada uma das celas, coberta com um capuz branco, que os obrigava a esconder a cara em todos os locais e em todos os momentos em que se pudessem

avistar mutuamente. José Luís Neto parte do negativo original da imagem, em formato 9 x 12 cm, precisamente identificado com o número “22474”, e procura nele fragmentos, isolando figuras, fotografando o grão e ampliando novamente, até tornar visível o interior de cada uma das celas, individualizando a mancha de cada ocupante. No fim realiza um projecto fotográfico com o retrato de cada um dos presidiários.

Figura 15 – Projecto 22474 de J.L.Neto e registo oficial de judeus em Portugal de Blaufuks

Daniel Blaufuks, que dedicou alguns dos seus últimos projectos à procura das suas raízes familiares judias, recorrendo à fotografia e ao vídeo. O seu último livro, chamado “Sob céus estranhos”, evoca a fuga dos refugiados judeus da Alemanha nazi e a sua passagem por Lisboa, a partir da história dos seus avós, também eles refugiados.

Temos por isso consciência de que o projecto fotográfico, que vamos realizar com base neste trabalho, se insere numa corrente ininterrupta de projectos que ao longo da história da fotografia têm utilizado o retrato e o auto-retrato como suporte.

(Almeida 1995); (Alves da Silva et al. 2006); (Barthes 1980); (Blaufuks 2007); (Fontcuberta 1988); (Friedlander 1970); (Frizot 1998); (Maddow 1977); (Michaels 1997); (Molder 1999); (Neto 2003); (Sander and Von Hartz 1991); (Sherman and Galassi 2003); (Vieitez 1998); (William 2006);

2.2 As teorias da percepção

De acordo com psicólogos e neurologistas, os sistemas de processamento visual humanos, à semelhança do que acontece com outros animais, dispõem de mecanismos complexos que parecem ter sido desenvolvidos para garantir um rápido reconhecimento de formas, em particular do rosto humano, e para possibilitar a aprendizagem a partir da experiência acumulada de percepções visuais. Estes mecanismos terão sido, como parece óbvio, essenciais para a sobrevivência das espécies que os desenvolveram e para o desenvolvimento de mecanismos sociais e culturais complexos. Todos nós temos consciência de que somos capazes de reconhecer um rosto num relance, ou mesmo de o adivinhar em menos de um relance.

O cérebro necessita de mecanismos optimizados para tratar de uma forma eficiente a quantidade enorme de informação que recebe e dela tirar partido para reconhecer formas, entre outras propriedades, e para fazer a sua aprendizagem.

Desde há muito tempo que psicólogos e neurologistas realçam um conjunto de evidências que indiciam a importância que as regularidades numa imagem têm na forma como a percepção funciona. Como exemplos de mecanismos mais básicos que evidenciam este tipo de fenómenos temos a adaptação do olho humano à luminosidade média, permitindo de uma forma automática a visão adaptada à iluminação, ou a adaptação também ao contraste de uma cena (variância), que parece ocorrer no córtice visual. Parecem existir evidências de adaptações a tipos mais complexos de regularidades como covariâncias entre entradas distintas.

(Attneave 1954); (Barlow 1989); (Barlow 2001); (Gross and Sergent 1992); (Gross 2002);

2.2.1 Primeiros passos

Attneave, na década de 50, terá sido o primeiro a sugerir a importância da redundância de informação nas imagens naturais e na forma como seriam apreendidas. Estas redundâncias possibilitariam uma representação mais económica, tirando partido por exemplo do conceito de limite ou fronteira, fazendo uma representação baseada sobretudo nas transições objecto / não objecto. Attneave realçou sobretudo o aspecto de economia, na percepção e na representação, que se consegue eliminando as redundâncias, que não trazem informação suplementar.

Attneave foi também o primeiro a utilizar na psicologia os conceitos da teoria da informação em canais de comunicação, de Shannon e Weaver, em particular os conceitos de capacidade de informação e de redundância. Cada canal tem um volume de informação que o caracteriza mas, muitas vezes, a quantidade de informação efectiva é inferior, devido a redundância, ou seja dados que apenas repetem outros dados e não trazem associada qualquer nova informação. A grande novidade introduzida por Attneave consistiu na possibilidade de introduzir na análise das teorias da percepção factores mensuráveis.

Como veremos a questão da economia voltou a adquirir importância na formulação de conceitos importantes como “Unsupervised Learning”, a que chamaremos aprendizagem não supervisionada, e “Sparse Coding”, a que chamaremos codificação esparsa. Estes conceitos tiverem uma grande visibilidade a partir dos trabalhos do psicólogo Horace Barlow.

(Attneave 1954); (Barlow 1989); (Barlow 2001); (Gross 2002); (Trehub 1997) ;

2.2.1.1 Conceitos Básicos de teoria da informação

A teoria da informação permitiu pela primeira vez quantificar e medir a informação associada a um determinado acontecimento. Essa medição assenta no conceito de probabilidade de ocorrência desse acontecimento. A quantidade de informação é calculada a partir do valor dessa probabilidade e é tanto maior quanto menor for essa probabilidade. A quantidade de informação é medida em bits.

Definição: 1 bit é a quantidade de informação que recebemos, quando ocorre 1 de 2 acontecimentos equiprováveis.

O valor que mede a quantidade de informação de um evento corresponde em princípio ao número de bits necessários para representar todas as hipóteses para o acontecimento em análise. De acordo com a teoria, a quantidade de informação de um acontecimento é dada por $\log_2 N$, onde *N* é o número de hipóteses possíveis para esse acontecimento, se todas essas hipóteses possíveis forem igualmente prováveis; $\log_2 N$ mede neste caso precisamente o número de bits necessários

para poder distinguir esses N valores. Por exemplo, a idade de uma pessoa, com valores entre 0 e 127 anos, vai necessitar no máximo de 7 bits para ser representada, pois os 7 bits permitem precisamente escrever números entre 0 e 127. Ao sabermos a idade de uma pessoa receberemos portanto uma quantidade de informação de 7 bits.

A teoria introduz ainda um conceito fundamental, designado por entropia. A entropia de uma variável aleatória X mede de alguma forma a quantidade de informação necessária para, em média, descrever a variável. O conceito de entropia tem em conta a probabilidade de ocorrência de cada estado de X, probabilidades que podem ser diferentes e não independentes. Se as probabilidades forem iguais verifica-se que a entropia é máxima. No exemplo acima das idades, a entropia máxima da variável idade é de 7 bits.

Um outro conceito da teoria de informação, fundamental para Barlow, na teoria que desenvolveu a partir do trabalho de Attneave, para explicar o processo de aprendizagem automática, é o conceito de mínima codificação, ou seja encontrar códigos optimizados, com um mínimo de bits, para uma determinada entropia, se esta não for máxima. Poder-se-ia por exemplo concluir, para o exemplo das idades, que talvez se pudesse codificá-las com menos de 7 bits, por as idades não serem igualmente prováveis.

De acordo com a teoria de Shannon é possível encontrar sempre códigos de comprimento (nº bits) variável – em vez de um código de comprimento fixo, com o valor da entropia máxima – com comprimentos médios iguais à entropia, ou quando muito ao valor da entropia +1

$$H \leq L < H + 1 \qquad (1)$$

Estes conceitos da teoria da informação são melhor explicados no anexo 10.3.

(Shannon 1948);

2.2.1.2 Aprendizagem não supervisionada

O conceito de aprendizagem não supervisionada é um conceito chave para Barlow. Como o o termo indica é uma aprendizagem feita de uma forma automática, sem orientação externa. Para Barlow a percepção de imagens ao longo da vida desempenha um papel essencial nessa aprendizagem não supervisionada (*Unsupervised Learning*). Mas a aprendizagem não é possível sem o confronto com a memória. A questão que Barlow se coloca é: "Qual a fonte do que parece ser um conhecimento organizado do que nos envolve, que indicia a existência de um mapa cognitivo (confirmado em inúmeras observações de muitos psicólogos e neurologistas), ou de um modelo dessa envolvência no nosso cérebro?" De acordo com Barlow, ainda que essa estrutura que modela o mundo possa ser parcialmente explicada pela genética, a sua instanciação só é possível com a incorporação da informação recebida, neste caso informação visual, na memória. Barlow acredita que a identificação de regularidades estatísticas ou de redundâncias na informação desempenham um papel essencial na construção desse modelo.

Tal como para Attneave a teoria da informação é crucial no desenvolvimento das suas conjecturas. Só a teoria da informação lhe permite compreender o paradoxo aparente de, sendo as redundâncias essenciais ao conhecimento e à aprendizagem, elas parecerem ser eliminadas pelos sistemas sensoriais e pelas redes neuronais que as tratam. Por isso ele defende que pode ser inconveniente utilizar a palavra redundância para referir essa enorme quantidade de informação que caracteriza qualquer imagem natural, pois essa redundância corresponde de facto a conhecimento previamente adquirido. Nós aprendemos à custa do reconhecimento dessa redundância. Por exemplo num rosto humano, quase toda a informação corresponderá a informação conhecida de outros rostos humanos. A rápida percepção e reconhecimento de um rosto implicam ser capaz de focar no que o

diferencia, para o reconhecer, ou para aprender, se for um rosto diferente. É a redundância ou seja o reconhecimento de regularidades nas imagens, que segundo Barlow, permite a aprendizagem automática. E esta aprendizagem exige conhecimento anterior.

De acordo com Barlow para ser possível a aprendizagem automática, esta deve permitir a representação separada de uma multiplicidade de alternativas e terá de conhecer a probabilidade de ocorrência normal de cada uma. A aprendizagem faz-se detectando relações entre eventos, entre padrões, e essa detecção só pode existir com a estimativa, ainda que inconsciente, da probabilidade de ocorrência dessas e de outras relações.

Por outro lado para se poder reconhecer algo novo e diferente, o modelo tem de permitir comparar o que vemos com o que acontece habitualmente. Só assim alguma coisa poderá ser reconhecida como nova, ou como diferente.

Como pode então o cérebro perceber que A1 implica B1? Para detectar a relação causa-efeito tem de se produzir o conhecimento de que o número de vezes em que B1 segue A1 é superior ao expectável, se não existisse essa relação. Para conhecer a probabilidade expectável, ou se mede directamente o número de vezes que B1 se segue a A1, e se detectam variações nesse valor, ou se estimam de alguma forma as probabilidades p(A1) e p(B1), probabilidades expectáveis para cada um dos acontecimentos, assumidos independentes. Esse conhecimento permitirá verificar se a ocorrência de B1 é superior ao expectável, quando A1 ocorre. Mas este cálculo será falacioso, se os A_s não corresponderem a componentes independentes. Nesse caso haverá redundância entre os A_s e isso significa que as probabilidades de ocorrência de cada caso estarão contaminadas.

Esta aprendizagem obriga a um mecanismo que permita inconscientemente a decomposição em componentes independentes, ou seja mais uma vez, eliminando a redundância. Se os componentes A_s e B_s forem independentes será fácil detectar se

$$p(B1|A1) >> p(A1)*p(B1) \qquad (2)$$

Conhecendo os p(A) e p(B) consegue-se calcular a probabilidade de ocorrência de todas as combinações, desde que se assegure a independência da decomposição. Portanto para sermos capazes de detectar coincidências suspeitas que poderão corresponder a relações, teremos de conhecer já as probabilidades de ocorrência dos componentes que pretendemos ser capazes de relacionar.

De acordo com Barlow tudo se passa como se dispuséssemos de um filtro negativo que compara cada mensagem com o que seria de esperar receber, obtendo-se dessa forma como resultado a diferença, baseada no conhecimento anterior, conhecimento que também é actualizado com essa nova mensagem. As teses de Barlow à luz do conhecimento actual pecam talvez por serem algo racionalistas e mecanicistas: o cérebro parece utilizar princípios de economia funcional e heurísticas próprias dos sistemas dinâmicos que não parecem suficientemente considerados nas ideias de Barlow.

Não será difícil perceber que há entre o modelo proposto por Barlow para o processo de "Unsupervised Learning" e outras teorias relevantes no campo da percepção, construídas a partir da observação do comportamento do cérebro em diversos animais, alguma complementaridade. É o caso da teoria que defende a existência de um mapa cognitivo no cérebro, com zonas relativamente delimitadas, ou mesmo o conceito de "ensemble" de células da avó[2], células pretensamente especializadas em reconhecer a imagem da avó, em rigor especializadas em reconhecer determinados rostos, como o rosto da mãe ou da avó.

A constatação, confirmada em numerosas experiências, de que existem zonas delimitadas de

2 A denominação "células da avó" resulta da adopção de uma imagem utilizada por Jerry Lettvin numa história de 1969.

células envolvidas no reconhecimento de rostos e de que essas células respondem a todos os estímulos visuais, faz com que muitos especialistas defendam a teoria de que a percepção e o reconhecimento visual envolvem um conjunto de grupos de células especializadas, cada um dando um contributo individual para definir um conjunto de atributos que permitem em conjunto fazer o reconhecimento. Este modelo é muitas vezes referido como codificação esparsa (*sparse coding*). Este modelo não está no entanto provado e outros autores, de que se pode destacar o trabalho de simulação de Arnold Trehub, defendem um modelo diferente para a codificação esparsa. Com base em simulações, utilizando um modelo de rede neuronal, Trehub propõe que existe uma célula apenas responsabilizada por cada identificação, apesar de todas as células responderem e que existe um mecanismo de filtragem com base nas intensidades dos sinais, para fazer o reconhecimento.

Como é evidente nem Barlow nem outros especialistas defendem que estes modelos correspondem em absoluto ao que acontece no cérebro em termos de percepção visual e de aprendizagem automática. São apenas modelos construídos com recurso à matemática, à estatística e à teoria da informação, que conseguem mimetizar comportamentos observados, e que permitem porventura ajudar a perceber melhor os mecanismos no cérebro, para podermos vir um dia a saber de facto como essas mensagens fluem, são guardadas e são acedidas, como memória, no cérebro.

(Attick and Redlich 1992); (Attneave 1954); (Barlow 1989); (Barlow 2001); (Edelman et al. 2001); (Edelman, Intrator, and Jacobson 2002); (Gross 2002); (Puga 2001); (Trehub 1997); (Young M. P. 1992);

2.2.2 Teorias recentes da percepção visual

Desde meados do século XX que existem muitos trabalhos teóricos e experimentais, realizados por psicólogos e por médicos, sobretudo da área da neurologia, dedicados à percepção visual e em particular à forma como o cérebro humano reconhece, memoriza e identifica rostos. Para além dos trabalhos de Barlow e Attneave, que já referimos porque são pilares fundamentais, por terem sido os primeiros a estabelecer a ponte entre a percepção e a teoria de informação, tentaremos nesta secção referir alguns trabalhos mais actuais nesta área, que podem ter relação com o nosso trabalho.

Têm sido realizados inúmeros trabalhos de investigação para tentar identificar as áreas do cérebro envolvidas nos processos de identificação e de reconhecimento de faces. Biólogos, neurologistas e psicólogos reconhecem a importância do processo de identificação e de reconhecimento de faces que admitem ter sido objecto de um longo processo de evolução, transformando-o num mecanismo complexo e fundamental para a integração social e para as práticas culturais humanas. Em particular a expressão de emoções envolve complexos mecanismos de mobilidade muscular que terão sido acompanhados com a sofisticação da representação da informação facial ao nível dos componentes cognitivos. Estes mecanismos permitem ao ser humano não só reconhecer os outros elementos da espécie, mas decifrar um conjunto de mensagens sociais e deduzir um conjunto de alertas fundamentais para a sua sobrevivência. É hoje consensual existirem no córtice visual áreas particularmente sensíveis a imagens de faces.

Figura 16 – Zonas do cérebro associadas ao reconhecimento de rostos
As zonas a tracejado correspondem à área envolvida no reconhecimento de rostos

Observou-se ainda que estas células não respondem a uma determinada face mas respondem em conjunto com padrões distintos a faces distintas. Este resultado parece estar em sintonia com a ideia de codificação esparsa e com técnicas optimizadas de decomposição das imagens em componentes. Verificou-se também que os dois hemisférios, esquerdo e direito, tratam a informação de face de forma diferente. Aparentemente o hemisfério direito parece reconhecer a face como um todo, enquanto o esquerdo reconhece a face como um conjunto de padrões. Por outro lado, se o reconhecimento das faces parece acontecer nas áreas visuais do córtice temporal, muitos casos clínicos e trabalhos experimentais parecem indicar que a interpretação e a atribuição de informação semântica envolvem estruturas mais profundas como a amígdala.

Como trabalhos determinantes nesta área começaremos por referir o trabalho de Vicky Bruce e Andy Young por terem sido os primeiros a propor um modelo completo para os processos cognitivos associados ao reconhecimento de faces. Este modelo apresentado na Figura 17 procura responder a todas as questões colocadas pelos trabalhos teóricos e experimentais na área da percepção visual, bem como pelos resultados observados em pacientes com vários tipos de perturbação no reconhecimento visual. Propõem um modelo com vários componentes funcionais que colaboram no processo de reconhecimento e posteriormente na atribuição de informação semântica a esse reconhecimento.

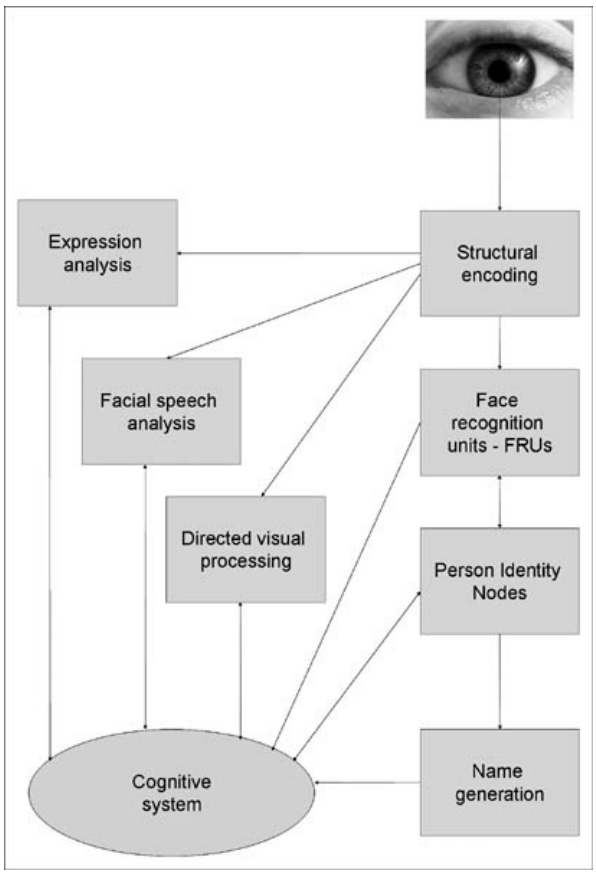


Figura 17 – Modelo Bruce Young

Estes autores consideram que existem sete níveis de codificação de informação associada a uma face humana: codificação pictórica, estrutural, semântica derivada do visual, semântica associada à identidade, nome, expressão facial e expressão de fala. Destas apenas a estrutural é utilizada no reconhecimento por FRUs, Unidades de Reconhecimento de Faces. Os autores defendem que o processo de reconhecimento é diferente para rostos familiares e não familiares. Isto deve-se ao facto de nos rostos familiares termos sido capazes de integrar informação estrutural ao longo de várias percepções e sermos por isso capazes de realçar os aspectos chave para o reconhecimento, ao contrário do que acontece com rostos novos, onde informação efémera pode ser determinante. Os autores descrevem e justificam o modelo, comparam-no e avaliam-no para outras áreas da percepção, como sejam o reconhecimento de objectos e de palavras escritas, realçando os pontos comuns e as diferenças.

Uma área a que os psicólogos tem dado particular atenção dentro do reconhecimento facial é a área da expressão de sentimentos e de emoções. Neste contexto teremos de realçar o trabalho de Paul Eckman que conduziu ao longo de muitos anos um trabalho experimental que, não só demonstrou a universalidade das expressões faciais, como definiu com precisão o modelo de movimentos musculares associados a cada expressão tipo.

Os resultados obtidos nestas áreas nos últimos anos conduziram ao aparecimento de um novo conceito agregador: as neurociências, que de alguma forma fazem a ponte entre a biologia, a neurologia e a psicologia, procurando integrar o comportamento biológico com os mecanismos sociais como a aprendizagem, a auto-regulação e todos os mecanismos desenvolvidos num contexto social. Estes trabalhos demonstraram a complexidade dos processos envolvidos, pois envolvem não só mecanismos fisiológicos, mas psicológicos e sociais, tal como o diagrama da Figura 18 evidencia.

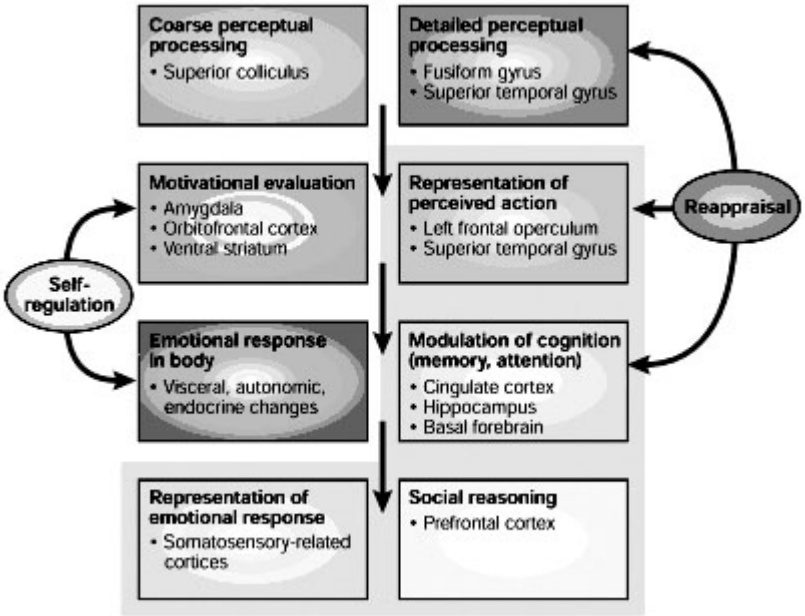


Figura 18 – Processos e partes do cérebro envolvidos na cognição social

Uma grande parte dos estudos nesta área parte de estímulos visuais, em particular a percepção de expressões ou de movimentos em rostos, como os representados na Figura 19.

Figura 19 – Retratos associados às expressão de emoções

De realçar o facto de existirem também diversos trabalhos experimentais e teóricos conduzidos nesta área na Faculdade de Psicologia do Porto, em particular no reconhecimento de expressões, com base numa BD de retratos, produzidos por um grupo de actores para esses projectos.

(Adolphs 2003); (Barbeau et al. 2008); (Barlow 1989); (Barlow 2001); (Bruce and Young 1986); (Ekman 1999); (Tovée 1995);

2.2.3 Teoria de detecção do sinal

A Teoria de Detecção do Sinal (SDT – *Signal Detection Theory*) é uma teoria de referência na área da psicologia, para analisar resultados de inquéritos relacionados com a percepção.

O modelo SDT oferece o suporte teórico para analisar resultados onde tenhamos de distinguir um sinal de um ruído perturbador.

O exemplo clássico para a instanciação do modelo SDT é oferecido por uma situação onde um controlador de radar tem de decidir, se aquilo que observa num ecrã, corresponde a um avião ou não. Se não existisse ruído capaz de produzir decisões erradas, este problema seria simples. Bastaria definir um limiar para tomar esta decisão. No entanto o ruído faz com que possam existir perante um contexto como o do operador de radar quatro situações possíveis.

Decisão	Avião presente	Avião ausente
Sim	ACERTO (HIT)	FALSO ALARME
Não	FALHA (MISS)	REJEIÇÂO CORRECTA

Tabela 1 – Os quatro tipos possíveis de resposta

A SDT oferece o suporte necessário para, a partir de dados estatísticos de decisões tomadas em situações que envolvem ruído, como a de um operador de radar, calcular dois parâmetros que quantificam e qualificam o processo. Para isso a SDT limita-se a assumir uma distribuição normal para o ruído e para o sinal[3]. Os dois parâmetros que quantificam o processo são **d'**, que mede de alguma forma a relação sinal/ruído e **ß** que mede o critério de decisão das pessoas envolvidas. É fácil perceber que as probabilidades obtidas para cada uma das quatro situações da tabela dependem por um lado da relação sinal/ruído – quanto maior mais acertos e menos erros – e do critério de decisão do operador – quanto mais conservador mais tendência terá para não identificar um avião, aumentando o número de falhas e de rejeições correctas, quanto mais atrevido maior o número de acertos, eventualmente identificando sempre os aviões, mas em contrapartida também aumentará o número de falsos alarmes.

Assumindo que o ruído é gaussiano e tomando, por simplificação, uma variável gaussiana normalizada, vamos ter um ruído com uma distribuição normal à volta de zero e um sinal+ruído tanto mais separado do ruído, quanto maior for a relação sinal/ruído, medida na SDT pela distância d', como se observa na Figura 20.

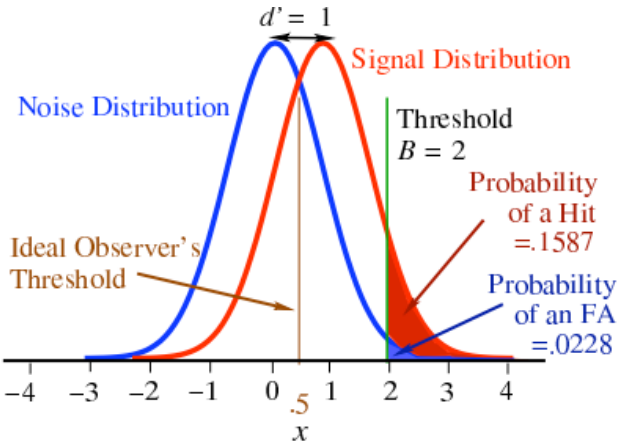


Figura 20 – Distribuições normais do ruído e do sinal na SDT

É a sobreposição parcial das duas curvas que dá lugar aos erros, aos falsos alarmes e às falhas. Se supusermos, por exemplo, que o critério de decisão do operador é o indicado por B=2, a probabilidade de um falso alarme será dada pela área a azul e a probabilidade de um acerto pela área a vermelho. Se as curvas se afastarem, como já tínhamos concluído, a relação sinal/ruído aumenta (d' aumenta) e a separação é mais fácil: a probabilidade de acertos aumenta e no limite será 1.

Tal como apontámos um modelo SDT é caracterizado pela variável **d'** e por uma variável que mede

3 Em rigor a SDT permite assumir outras distribuições, como veremos. No entanto a distribuição normal é aquela que dá melhores resultados em grande parte das situações.

o critério de decisão. Para esta última existem várias soluções, ainda que todas equivalentes: **B** – a distância do limiar de decisão a zero; **D** – distância do limiar de decisão ao sinal; **C** – a distância do limiar de decisão ao ponto médio ente o ruído e o sinal, ponto designado como observador ideal (valor para o qual a probabilidade de uma falha é igual à de um falso alarme); **ß** – relação entre os valores das duas funções de densidade de probabilidade, no nosso exemplo P(1)/P(2). Todos os parâmetros apresentados são equivalentes – a partir de um calculam-se facilmente os outros – mas cada um, por si, dá-nos informações diferentes. É fácil mostrar por exemplo que

$$\ln ß = d' * C \qquad (3)$$

B ou **D** permitem porventura uma melhor visualização gráfica na Figura 20, mas não nos dão qualquer indicação sobre as características de observador, pois a distância do ponto de decisão a zero ou ao sinal, por si, nada dizem sobre se o observador tem mais tendência a dizer sim ou não. Para ter esta perspectiva pode-se utilizar o parâmetro **C**, mas a maioria dos autores prefere o parâmetro **ß.** É fácil observar que se **C** é positivo o observador tem tendência para o não (conservador) e se é negativo tem tendência para o sim (liberal). Esta tendência será tanto maior quanto maior o valor absoluto de **C**. No caso do parâmetro **ß**, se o valor for maior do que 1 haverá tendência para o não, se for menor do que 1 tendência para o sim. O **ln ß** tem a vantagem de garantir uma distribuição linear dos valores entre -∞ e +∞: valores positivos para conservadores, negativos para liberais.

Conhecendo as 2 curvas gaussianas e o limiar de decisão é fácil calcular todas as probabilidades. Por exemplo a probabilidade de falha seria neste exemplo de 1 – 0.1587, ou seja a área limitada pela curva do sinal menos a zona a vermelho.

No entanto normalmente as curvas não são conhecidas. Pretende-se sim estimar os parâmetros e traçar as curvas a partir de um conjunto de dados estatísticos obtidos experimentalmente.

É fácil observar que as probabilidades para as quatro situações identificadas não são independentes. De facto

$$P(HIT)=1 - P(FALHA) \qquad (4)$$
$$P(FA) =1 - P(REJEIÇÃO\ CORRECTA) \qquad (5)$$

Bastará portanto conhecer a probabilidade de ACERTOS (HIT) e de FALSOS ALARMES (FA).

Podemos calcular todos os parâmetros do modelo a partir apenas de P(HIT) e de P(FA).

Consideremos a função conhecida e tabelada Z(P), que é a função que nos dá o valor de Z para o qual $P(x<Z)=P$. No fundo a função Z(P) calcula o valor de Z para o qual a área limitada pela função densidade de probabilidade, para $x < Z$, é igual a P, como se observa na Figura 21, onde o valor P corresponde à área delimitada pela função e marcada a amarelo.

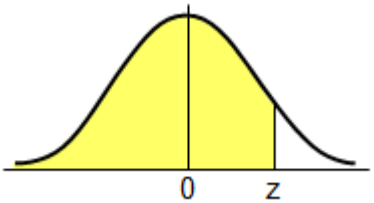


Figura 21 – A função Z(P) mede a área a amarelo

É fácil demonstrar que

$$d'=Z(P(HIT))-Z(P(FA)) \qquad (6)$$

$$B=Z(1-P(FA)) \quad (7)$$

$$C=-1/2(Z(P(HIT))+Z(P(FA)) \quad (8)$$

$$\ln \beta = d'(B-1/2d') \quad (9)$$

O modelo pode também ser ilustrado por curvas que mostram a relação entre a P(HIT) e a P(FA), para valores constantes de d', tal como se ilustra na Figura 22. São as chamadas curvas ROC (Receiver Operating Carachteristics) ou curvas de isosensibilidade, pois cada curva corresponde a manter constante a relação sinal ruído (d´), fazendo variar o critério de decisão. Cada ponto de cada curva ROC corresponderá a um valor do critério de decisão.

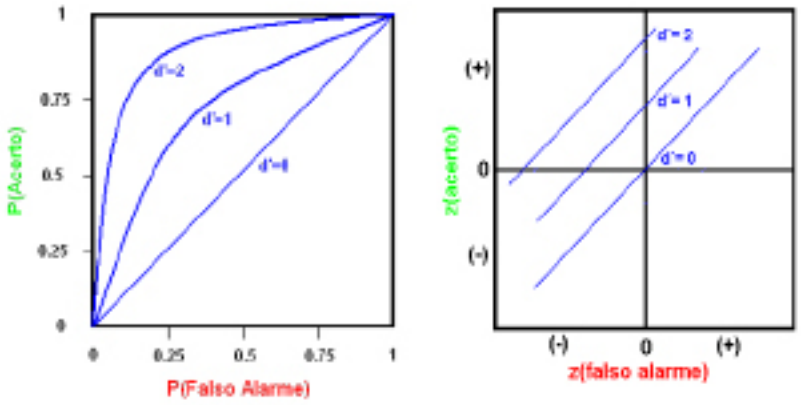


Figura 22 – As curvas ROC para vários valores de d'

Estas curvas poderão ser traçadas utilizando coordenadas probabilísticas (P(HIT) e P(FA)), ou utilizando coordenadas gaussianas (Z(P(HIT)) e Z(P(FA))), transformando as curvas ROC da 1ª figura nas linhas paralelas representadas na 2ª figura. Pode-se facilmente concluir que um determinado valor do critério de decisão **B** corresponde a traçar uma linha vertical na figura. Qualquer ponto numa ou noutra figura determina portanto o valor de **d'**, pela curva ROC onde se situa, e o critério de decisão **B**, pelo valor da abcissa.

(Abdi 1987); (Abdi 2007); (Harvey 2009); (Heeger 2003); (Wickens 2001);

2.2.3.1 Avaliação do modelo

Ao recorrer ao modelo SDT um dos problemas que se coloca é de avaliar se o modelo foi capaz de traduzir os resultados da experiência realizada. Uma das formas comummente aceite de o fazer é recorrer à distribuição *chi-square*. Esta distribuição caracteriza variáveis aleatórias Z que resultam da soma dos quadrados de variáveis gaussianas

$$Z = X^2 = X_1^2 + X_2^2 + ...X_n^2 \quad (10)$$

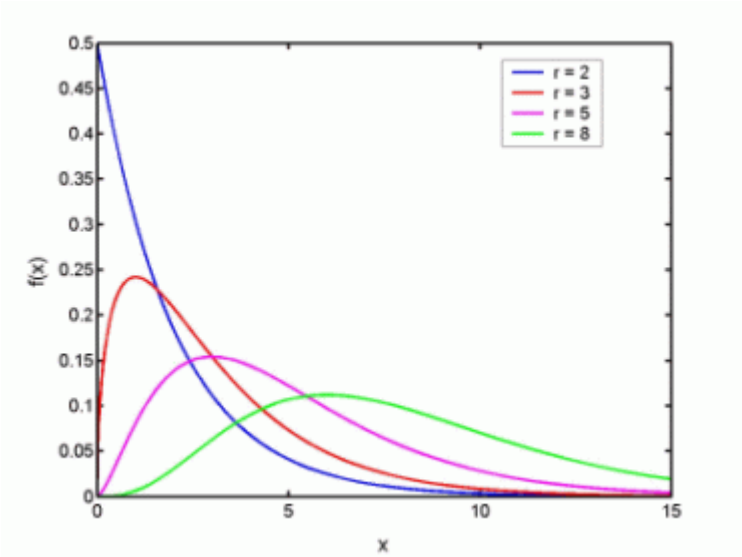


Figura 23 – A distribuição chi-square para vários valores do número de graus de liberdade

Ao contrário da distribuição normal que tem sempre a mesma forma, onde apenas a média e a variância podem mudar, verifica-se que na distribuição chi-square a forma da curva depende de um parâmetro, o número de graus de liberdade (r), que determina também a média e o desvio padrão, respectivamente r e 2r. Aumentando r a curva tende para a distribuição gaussiana, aliás como seria de esperar se tivermos em conta o teorema do limite central.

Para validar o modelo SDT analisa-se o valor de uma função, designada como estatística de Pearson

$$X^2 = \sum_{i=1}^{c} (x_i - m_i)^2 / m_i \qquad (11)$$

onde cada m_i é o valor estimado pelo modelo para cada x_i.

Como é evidente o modelo será tanto melhor quanto menor for a estatística de Pearson, pois tal significa que os valores estimados e observados são próximos. Idealmente seria zero. Pode-se demonstrar que a estatística de Pearson, dentro de determinados condicionalismos, tem uma distribuição chi-square. Ora esta distribuição é bem conhecida e podem-se facilmente calcular os valor da probabilidade associada a $X^2 > A$.

Concluindo, em função do número de graus de liberdade, podemos afirmar que o modelo testado é bom, se o valor da probabilidade obtido pela distribuição chi-square, para o valor de X^2, calculado pela expressão (11), for significativamente superior a 0.5, pois isso garante que a probabilidade de o resultado ser fruto do acaso é bastante pequena. Por exemplo, para r=8, se o valor de X^2, calculado por (11), for bastante menor que 7.34, podemos aceitar a validade do modelo - pois para r=8 p(7.34)=0.5.

(Wickens 2001);

2.2.3.2 Generalização do modelo

A teoria foi inicialmente desenvolvida para situações, como a descrita, onde há a presença de um sinal e de um ruído, que perturba a leitura desse sinal, mas tem sido adoptada em inúmeras

situações, nomeadamente na área da psicologia e da sociologia, para avaliar os resultados de inquéritos, bastando para tal alargar o conceito de sinal e de ruído. É o caso por exemplo de situações relacionadas com o reconhecimento visual de imagens. Neste caso pode considerar-se que o sinal é o sinal correspondente ao reconhecimento feito no córtice visual, quando se vê uma imagem conhecida, e o ruído o sinal que aparece na sua ausência.

A teoria genérica, apresentada com um sinal e com um ruído, ambos com distribuições normais, pode também ser facilmente generalizada para contemplar:

- situações de vários níveis de sinal: S1, S2,… Sn
- situações com vários níveis de critério de decisão: em vez de um simples Sim/Não poderemos ter por exemplo Não/ Sim com pouca certeza/ Sim com certeza razoável/ Sim com certeza absoluta
- situações onde o desvio padrão do sinal não é igual ao do ruído
- situações com outras distribuições de probabilidade, não gaussianas
- situações onde em vez de um sinal e de um ruído temos um vector de sinais e um vector de ruídos, ou seja temos variáveis aleatórias de dimensão N

Todas as expressões apresentadas nas secções anteriores podem ser facilmente generalizadas para cada uma destas situações.

Dada a pertinência deste modelo para os resultados obtidos neste trabalho com base em inquéritos sobre o reconhecimento de retratos reconstruídos com erro, utilizando as técnicas estatísticas que experimentámos, decidimos validá-los utilizando a Teoria de Detecção de Sinal.

Na secção 6 apresentamos o modelo que utilizámos e analisamos os resultados.

(Wickens 2001);

2.3 Modelos de computação e processamento de imagem

Partindo de teorias da percepção, onde as propostas de Barlow têm um papel relevante, houve, a partir dos anos 80, um trabalho continuado no domínio da computação e processamento de imagem para construir modelos para tratar imagens naturais. No caso das imagens de rostos as motivações foram diversificadas, mas temos de reconhecer que talvez a mais importante tivesse sido e continue a ser a possibilidade de reconhecer a identidade de utilizadores de diversas aplicações, com vantagens óbvias, nomeadamente na área da segurança, onde ainda hoje somos obrigados a memorizar um sem número de logins e passwords para entrar no nosso PC, ou para nos ligarmos à rede, e para aceder a todo o tipo de aplicações locais ou na internet.

Existem dois grandes tipos de técnicas para reconhecer uma imagem de um rosto.

O primeiro grupo faz essa identificação baseada na medição de um conjunto de parâmetros-chave. São chamadas por isso técnicas paramétricas (*feature based*). A grande vantagem destas técnicas reside na independência das condições de iluminação, que constitui um dos factores mais difíceis de ultrapassar nas técnicas estatísticas. São no entanto extremamente dependentes do rigor de medição dos valores paramétricos utilizados, o que na maioria dos casos é extremamente difícil de garantir, sobretudo se pretendermos aplicações, onde a cooperação da pessoa a identificar não é obrigatória. São, por isso, técnicas com grau de falibilidade elevado.

O segundo grupo faz uma identificação global. Neste caso, os valores que obtêm e que servem para fazer a identificação não estão de forma nenhuma relacionados com zonas específicas do rosto, nem com parâmetros físicos, não podendo por isso ser associadas a qualquer característica visível. São chamadas técnicas estatísticas, por se basearem em análises estatísticas de conjuntos de rostos humanos.

Neste caso, as imagens são tratadas como vectores de dimensão N, onde N representa o número de pixéis.

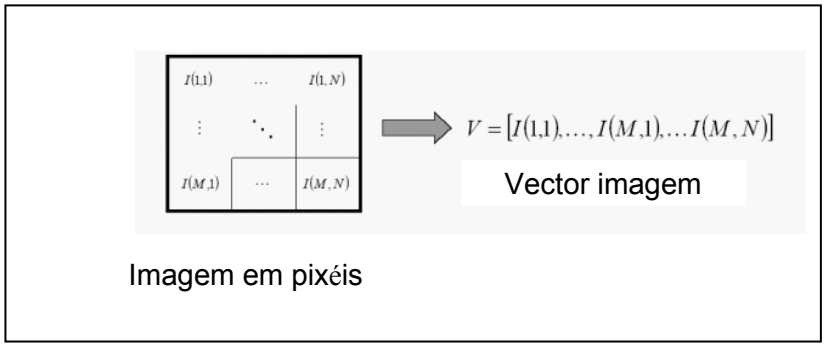


Figura 24 – Imagem e vector imagem

Utilizando técnicas estatísticas sobre um conjunto de vectores imagem, obtidos por serialização de n imagens de rostos, extraem-se por métodos estatísticos um conjunto de vectores componentes, que vão corresponder a uma BD de conhecimento de imagens.

Se identificarmos as imagens a tratar como vectores, num sistema dimensional de ordem N, onde N é o número de pixéis de cada uma, ou seja o número de linhas x número de colunas, a operação permite passar de um sistema dimensional de ordem elevadíssima N, para um sistema dimensional de ordem muito mais baixa, onde a nova ordem corresponde ao número de componentes, tipicamente inferior ao número de imagens que constitui a BD de imagens que se utiliza. Cada vector original pode então ser reconstituído, somando as projecções nos K (K << N) vectores que foram identificados como componentes.

Existem diversas técnicas / algoritmos que utilizam a decomposição das imagens em componentes. É o caso do PCA (Principal Component Analysis), do SCA (Symetric Component Analysis), do ICA (Independent Component Analysis), do LDA (Linear Discriminant Analysis), ou do NMF (Non Negative Matrix Factorization). Pertencem todos ao segundo grupo de técnicas não paramétricas.

O PCA, o ICA e o NMF, que vão ser as técnicas utilizadas neste trabalho, vão ser descritas com algum pormenor nas secções 2.3.4, 2.3.5 e 2.3.6.

O LDA, por parecer interessante para futuros desenvolvimentos deste trabalho, ao introduzir o conceito de classe nas imagens, vai também ser apresentado sumariamente na secção 2.3.7.

Qualquer uma das técnicas referidas permite decompor qualquer imagem natural numa combinação linear de componentes, como se tivéssemos disponível um dicionário ou um alfabeto imagético, para depois escrever com ele qualquer imagem.

Todos os métodos estatísticos assumem que a partir de uma imagem I, ela pode ser vectorizada e depois pode ser desdobrada numa constante e num somatório de componentes I_k, com pesos diferenciados, dados por coeficientes α_k, utilizando cada método critérios e algoritmos diferentes, para encontrar esses componentes e fazer essa decomposição.

$$I = I_{\varnothing} + \sum_k a_k I_k \qquad (12)$$

Estes modelos revelaram-se poderosos, nomeadamente em aplicações de reconhecimento de imagem, por exemplo para reconhecer a identidade de alguém a partir da imagem de um rosto,

utilizando como suporte para a identificação uma BD de atributos de imagens.

Um sistema de reconhecimento de uma imagem, em particular de um rosto, consiste habitualmente de quatro grandes subsistemas, identificados na Figura 25.

A **detecção de face** reconhece a região da imagem, onde a face a identificar se localiza. Este subsistema é particularmente importante e complexo em sistema dinâmicos, como por exemplo vídeos, onde a face se movimenta.

O **pré-processamento da face** normaliza a imagem da face para garantir que a extracção dos parâmetros físicos ou estatísticos é eficiente. Esta normalização pode contemplar um conjunto de operações desde alinhamentos (translações, rotações), mudança de resolução, escalas, etc.

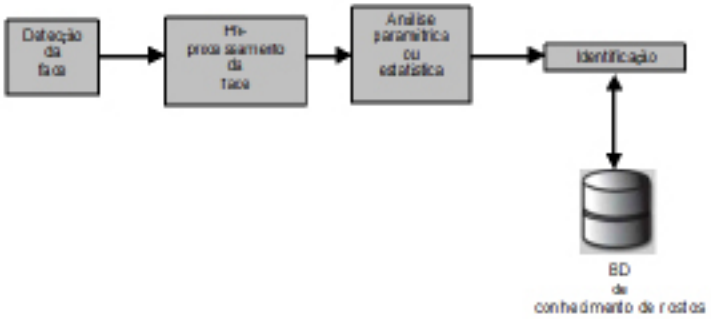


Figura 25 – Sistema de reconhecimento de rosto

A **análise** analisa a imagem para obter o conjunto dos parâmetros a utilizar na identificação.

Por fim a **identificação** compara os parâmetros da imagem com os valores conhecidos da BD, para concluir se se trata de um rosto conhecido ou não.

Como é evidente ninguém pretende afirmar que nenhum dos modelos desenvolvidos reflectem em absoluto o comportamento do cérebro na percepção de imagens, na sua identificação ou no processo de aprendizagem automática, mas para além de se terem revelado importantes para aplicações que, de alguma forma pretendem substituir os humanos em tarefas de percepção e de aprendizagem visual, podem constituir indicadores importantes sobre a forma como o cérebro e os neurónios dedicados à percepção poderão optimizar o processo e tirar partido, na identificação e na aprendizagem, da redundância da informação.

(Bartlett 1998); (Gross, Shi, and Cohn 2001); (Xudong 2006); (Yang 2002); (Yun 2007); (Zhao et al. 2003);

2.3.1 A evolução das técnicas estatísticas

O desenvolvimento desta área acontece na década de noventa com o aparecimento de um conjunto de técnicas e de algoritmos que permitem fazer a decomposição de imagem de rosto para desenvolver aplicações, vocacionadas para o reconhecimento de identidades, em várias áreas, desde o entretenimento até à vigilância e à segurança no acesso a informação. Na área da segurança a ideia chave é a de criar soluções amigáveis e muito seguras, que substituam memorizações de PINS ou de *passwords*, para aceder a informação informatizada.

Surgem revistas especializadas, conferências de renome, como a AFGR (*Automatic Face and Gesture Recognition*) do IEEE ou a AVBPA (*Audio and Video based Biometric Person Authentication*) e também um conjunto de Bases de Dados de retratos, utilizadas em vários projectos de investigação.

Depois de alguns anos em que surgiram alguns trabalhos nesta área e as primeiras aplicações, foi desenvolvido um esforço para conseguir normalizar protocolos de avaliação das técnicas/ algoritmos/aplicações, permitindo definir métodos globalmente aceites para comparar a sua performance, de uma forma estruturada e em diferentes situações bem definidas. Este trabalho foi desenvolvido, associado a uma Base de Dados muito conhecida, pois constituiu a base de muitos projectos de investigação nesta área, a Base de Dados de FERET. O protocolo de avaliação ficou assim conhecido como o protocolo de FERET, protocolo que foi utilizado para diferentes testes ao longo da década de 90.

Estes testes permitiram avaliar sucessivamente algoritmos mais poderosos, observando que no entanto nem sempre os mesmos algoritmos davam os melhores resultados em todas as situações. Muitos destes algoritmos foram melhorados precisamente com base nos resultados dos testes de FERET. Constatou-se por outro lado que muitas vezes é difícil separar os resultados associados à técnica de representação (por exemplo PCA ou ICA) da forma de implementação dessa técnica.

Com base no protocolo de FERET foi desenvolvido já no ano 2000 o FRVT (Face Recognition Vendor Test), que tem vindo a ser utilizado em novos testes a partir desse ano.

Muitos destes testes têm sido conduzidos para avaliar a performance das técnicas/ algoritmos em situações diversas, que tipicamente dificultam o reconhecimento, tais como:

- retratos com diferentes ângulos de visão
- retratos com diferentes condições de iluminação
- retratos com diferentes expressões faciais
- retratos com diferentes áreas de oclusão
- retratos com intervalos de tempo significativos
- retratos em movimento (vídeo)
- outras

Estes testes têm sido conduzidos utilizando Bases de Dados especialmente vocacionadas para cada condição a testar.

Podemos realçar a avaliação feita pela CSU (Colorado State University) a vários algoritmos, utilizando a BD de Feret, para o que desenvolveu um conjunto de aplicações *open-source* para realizar os testes especificados.

Constatamos que um número significativo de trabalhos de investigação e de teses de dissertação continuam nos últimos anos a focar na obtenção de algoritmos mais eficazes e mais eficientes, em particular vocacionados para responder com sucesso a situações difíceis, como sejam variações significativas na iluminação, ou a oclusão parcial do rosto, ou a reconhecer e identificar expressões diversificadas. Quase sempre estes novos algoritmos correspondem a variações ou a melhorias introduzidas nos métodos de base, como sejam o PCA ou o ICA. Mas aparecem também propostas de métodos diferentes. É o caso por exemplo do NMF, método recente que o nosso trabalho também pretende abordar, onde diversas teses apareceram, sobretudo a propor algoritmos diferentes, porventura mais eficientes, para conseguir fazer a decomposição NMF, e comparando os resultados obtidos como o NMF com outras técnicas, como o PCA e o ICA.

O nosso trabalho não vai estar centrado na melhoria ou na optimização de nenhum dos algoritmos utilizados no reconhecimento de rostos, não procurando portanto qualquer inovação nessa área. Vai limitar-se a utilizar alguns deles para calcular os componentes e construir as diversas BD de conhecimento de rostos que irão suportar o nosso trabalho. Desenvolvemos depois um conjunto de scripts suportadas em MATLAB para validar e comparar resultados no que diz respeito à capacidade

de reconstrução de retratos, em particular com poucos componentes, testando também modelos híbridos, que associam diferentes tipos de componentes calculados por diferentes algoritmos.

(Bolme et al. 2003); (Gross, Shi, and Cohn 2001); (Xudong 2006); (Yang 2002); (Yun 2007); (Zhao et al. 2003);

2.3.2 A relação com a psicologia e a neurologia

Tem havido ao longo do tempo uma proximidade entre os projectos de computação e processamento de imagem nesta área e o conhecimento adquirido em áreas como a psicologia e a neurologia. Muitos conhecimentos e teorias propostas por psicólogos e neurologistas têm influenciado algoritmos ou técnicas utilizadas em sistemas de reconhecimento.

Sabe-se bem, por um lado, que será muito difícil, para não dizer impossível, no futuro próximo, construir um sistema com o poder e a subtileza dos sistemas de percepção humanos, mas acredita-se que o progresso nesta área poderá ajudar a compreender a percepção humana. Por outro lado a capacidade de trabalhar com milhões, biliões de imagens, que um sistema informático oferece, ultrapassa largamente os limites impostos em dimensão pelo cérebro humano.

Existe de facto um conjunto de perguntas que têm vindo a ser colocadas na área da psicologia e da neurologia que de alguma forma têm constituído vectores para o trabalho nesta área. Perguntas como:

- Existe um sistema dedicado para o reconhecimento de faces (as designadas células-avó) – Um conjunto de estudos recentes parece confirmar esta ideia de que existe no sistema de percepção visual humano um conjunto de mecanismos optimizados para o reconhecimento e para a aprendizagem de rostos humanos. Este vai ser também um dos vectores do nosso trabalho, pois pretendemos contribuir para validar a tese de que o sistema de percepção, de aprendizagem e de reconhecimento pode ser optimizado à custa de mecanismos como os PCAs ou ICAs, permitindo um reconhecimento simples, rápido e com um consumo energético mínimo, condições que parecem ter sido essenciais para inclusive a sobrevivência da espécie.
- A percepção de faces é orientada a valores paramétricos ou é antes global – Curiosamente existem também disponíveis, como já referimos, técnicas e algoritmos de reconhecimento computacional orientadas precisamente a valores paramétricos, por exemplo algumas medidas chave que permitam fazer a identificação e técnicas e algoritmos de reconhecimento orientadas à percepção global, baseadas em técnicas estatísticas. É o caso das técnicas que utilizam PCA, SCA, ICA ou LDA. Serão as técnicas que iremos abordar neste trabalho.
- Quais os elementos da face mais importantes para um reconhecimento – olhos, boca, nariz, orelhas? A parte superior ou a parte inferior da face? Este problema pode ser importante mesmo em técnicas estatísticas para possibilitar reduzir significativamente a dimensão das imagens a tratar. O recurso ao ICA ou NMF, por utilizar factores que parecem corresponder a zonas do rosto, pode ajudar a obter alguns dados que suportem uma resposta a esta questão.
- Existe relação entre algumas características de determinadas faces e facilidade de reconhecimento/aprendizagem. Este aspecto não será abordado pelo nosso trabalho.
- Efeitos resultantes de diferentes expressões faciais, condições de iluminação, oclusão, passagem de tempo, etc. – muitos estudos se têm preocupado com estas questões, como referimos, nomeadamente porque parecem ser importantes em aplicações na área da segurança. Estes aspectos também não serão abordados pelo nosso trabalho.

(Barlow 1989); (Barlow 2001); (Gross 2002); (Zhao et al. 2003)

2.3.3 Técnicas estatísticas

Nesta secção apresentaremos com algum detalhe as técnicas estatísticas que utilizámos no nosso trabalho.

Para utilizar estas técnicas teremos de começar por seriar cada imagem a tratar, num vector de dimensão N, onde N é o número de pixéis. Obtemos, como resultado da vectorização de n imagens,

uma variável vectorial X, representada por uma matriz, em que cada coluna corresponde a uma observação.

$$\begin{matrix} x_{11} & x_{12} \ldots\ldots x_{1n} \\ x_{21} & x_{22} \ldots\ldots x_{2n} \\ \ldots & \\ x_{N1} & x_{N2} \ldots\ldots x_{Nn} \end{matrix} \qquad (13)$$

Obtemos assim n observações, em que cada observação corresponde a uma coluna da matriz X, um vector de comprimento N (onde tipicamente n << N). Todas as técnicas tentam decompor essa matriz de imagens vectoriais X, utilizando um determinado critério nessa decomposição. O problema genérico reside em encontrar uma função f(X) que garanta que a matriz U, transformada de X, satisfaz esse critério.

$$U=f(X) \qquad (14)$$

As decomposições que vamos analisar utilizam todas transformações lineares, ou seja

$$U^T = W*X^T \qquad (15)$$

Na equação (15) U^T e X^T são simplesmente as matrizes U e X transpostas, utilizadas em vez de U e X por facilidade de cálculo. Nas matrizes transpostas as imagens correspondem a linhas em vez de colunas.

Portanto o problema traduz-se em calcular a matriz W, que permita obter a transformação f() pretendida, baseada nos critérios definidos, e encontrar os componentes desejados, como a Figura 26 a representa. Conhecendo W, podemos calcular a matriz inversa de W, a matriz A, por vezes designada como matriz de mistura e obtemos a equação (16).

$$X^T = A*U^T \qquad (16)$$

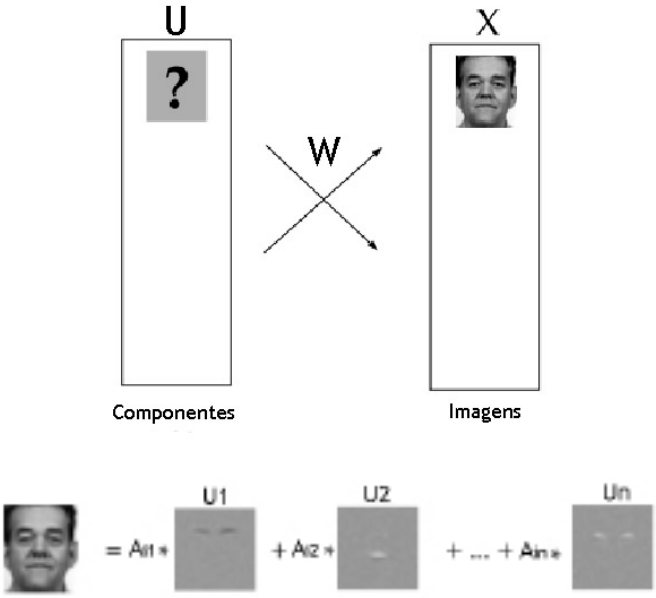


Figura 26 a e b – Reconstrução de imagens

A Figura 26 b ilustra a equação (16): as imagens podem ser reconstruídas por uma mistura pesada (coeficientes A_{ij}) dos componentes U. Visto de outra forma, todas estas técnicas permitem encontrar para uma variável vectorial X (em que cada instância da variável se pode associar a

um ponto num espaço referencial de N dimensões) um novo sistema de eixos $U(u_1, u_2,\ldots u_n)$, que permitem decompor cada imagem (vector Xi) numa série de componentes, construídos sobre cada um desses novos eixos.

Portanto todos os métodos estatísticos factorizam um conjunto de imagens em componentes, que são obtidos utilizando critérios / restrições diferentes, por exemplo no caso do ICA, a que corresponde a imagem da figura, devemos obter novas variáveis que devem ser estatisticamente independentes, ou seja, dado um valor para uma variável num componente nada se pode afirmar sobre outras variáveis noutros componentes.

Se nos atrevermos a fazer uma analogia com o modelo neuronal, o cálculo de cada componente poderia corresponder a um conjunto de células, a que corresponderia, para cada imagem, um peso A_k, que esse conjunto teria a capacidade de actualizar por uma aprendizagem não supervisionada.

Diversos métodos, baseados em critérios diferentes, têm sido propostos, para transformar a matriz X. Sendo o critério diferente e norteado também por objectivos diferentes – que pode ser simplesmente a diminuição da dimensão dos dados, ou encontrar componentes significativos sobre qualquer aspecto especial, ou até um critério muito orientado para uma aplicação específica, não se poderá afirmar que um é melhor do que outro, pois todos têm vantagens e desvantagens e são mais indicados para determinadas aplicações.

Nota: os algoritmos que realizam técnicas estatísticas como o PCA e o ICA fazem sempre um pré-processamento dos dados. Este pré-processamento inclui a subtracção das médias aos dados, para obter dados centrados em zero. Em vez da matriz de dados X trabalharemos com uma matriz X_E.

Cada vector imagem começa por ser normalizado, tomando como referência a média desse vector. Define-se a média do vector X_i como $E[X_i]$

$$E[X_i] = 1/N \sum_{i=1}^{N} X_j \qquad (17)$$

Teremos portanto uma matriz X com vectores Xi normalizados - ver equação (18) - ou seja com médias iguais a zero.

$$X_i = X_i - E[X_i] \qquad (18)$$

Antes de fazer o cálculo é ainda subtraído o vector médio da matriz X, garantindo assim valores médios zero para as médias dos valores de cada pixel.

$$X_E = X - E[X] \qquad (19)$$

O objectivo é de apenas de diminuir a complexidade dos cálculos, em particular da matriz covariância, por se trabalhar com variáveis com média zero, pois é simples depois reconstituir os resultados, tendo em conta o valor médio retirado.

Iremos por isso assumir sempre, nas expressões do PCA e do ICA, dados pré-tratados e centrados em zero. Já no caso do NMF as médias, como veremos, não podem ser retiradas, em virtude do NMF impor valores positivos.

(Bartlett 1998); (Hyvärinen 1999, 1999); (Hyvärinen and Oja 2000); (Hyvärinen, Oja, and Juha 2001); (Xudong 2006); (Yun 2007);

2.3.4 Principal Component Analysis - PCA

Os primeiros autores a propor e a popularizar o PCA para tratar imagens de rostos foram Kirby, Sirovich, Turk e Pentland. O PCA (*Principal Component Analysis*) impõe como critério da função de transformação f() – ver equação (14) - que os componentes correspondam às direcções de variância máxima da informação e sejam ortogonais entre si, ou seja não correlacionados. Este critério é o indicado quando se pretende diminuir a dimensão dos dados e minimizar o erro da decomposição. Por isso, a direcção do primeiro componente será calculada como a direcção onde a variância das projecções dos vectores X_i é máxima. A Figura 27 permite observar precisamente esse facto.

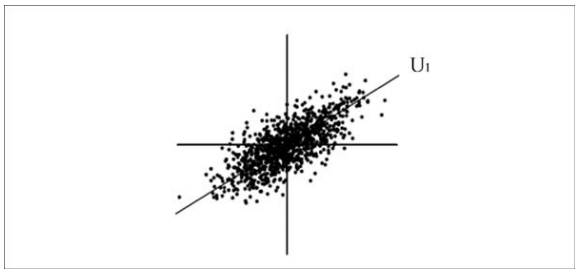


Figura 27 – 1º eixo PCA para dados 2D

No conjunto de dados 2D representados na Figura 27, o 1º eixo PCA seria o eixo U_1. É fácil perceber que se pretendêssemos substituir cada um dos vectores 2D, associados a cada ponto, por um só componente, esta solução minimizaria o erro.

As técnicas de PCA permitem portanto encontrar, para uma variável vectorial X (em que cada instância da variável – cada coluna da matriz X - se pode associar a um ponto num espaço referencial de N dimensões), um novo sistema de eixos ortogonais, que permitem decompor esses vectores numa série de componentes, construídos sobre cada um desses eixos.

Como simples ilustração desta técnica poder-se-ia recorrer à decomposição de um sinal numa série de Fourier, onde cada componente da série corresponde de facto a um coeficiente num eixo de um sistema de eixos ortogonais.

(Jackson 1991); (Jolliffe 2002);

2.3.4.1 Como implementar o PCA?

Depois de ler a matriz X com todos os vectores imagem faz-se a pré-normalização dos dados, descrita nas equações (17), (18) e (19).
Ao fazer a decomposição PCA, a partir da matriz vectorial X normalizada, cada vector Xi vai ser decomposto em componentes que correspondem à projecção desse vector em cada um dos eixos do novo sistema ortogonal encontrado. O objectivo do método PCA é obter um sistema ortogonal de vectores U, que caracterizem tão bem como possível a distribuição dos vectores iniciais. No PCA procura-se um novo sistema de vectores que maximize as variâncias das projecções dos vectores imagens. Vamos procurar portanto vectores U_k que coincidam com as direcções de máxima variância das imagens. Feita esta decomposição, o vector imagem de ordem i, X_i, pode ser reconstruído com um conjunto de coeficientes A_i, somando, de uma forma pesada, os componentes U_K, de acordo com a equação (20).

$$X^T_i = Ai*U^T \qquad (20)$$

A matriz U contém os vectores que identificam os eixos do novo sistema. Nessa matriz cada coluna U_k vai conter os valores do vector que identifica cada um dos eixos no novo sistema de eixos.

$$\begin{matrix} u_{11} & u_{12} \ldots\ldots u_{1n} \\ u_{21} & u_{22} \ldots\ldots u_{2n} \\ \ldots & \\ u_{N1} & u_{n2} \ldots\ldots u_{Nn} \end{matrix} \qquad (21)$$

Poder-se-ia afirmar que cada componente PCA corresponde a um atributo global (*feature*) da imagem neste modelo de representação e que cada coeficiente mede esse atributo. Esta designação é no entanto polémica pois alguns autores preferem reservar a palavra '*feature*' para propriedades físicas, identificáveis.

Para encontrar os vectores U_k que coincidem com as direcções de máxima variância das imagens teremos de maximizar cada Δ_k

$$\Delta_k = 1/N \sum_n (U_k^T * X_n)^2 \qquad (22)$$

o que significa maximizar o somatório dos módulos das projecções dos vectores X nos novos eixos, onde U_k é o componente PCA de ordem k e a ortogonalidade dos U_k impõe:

$$U_i^T * U_K = 1 \text{ ou } 0 \text{ conforme } i=k \text{ ou } i!=k \qquad (23)$$

Baseado na teoria das matrizes, os vectores U_k serão os vectores próprios (*eigenvectors*) da covariância de X. Calcule-se então a matriz C das covariâncias de X.

$$C = cov(X) = 1/n(X*X^T) \qquad (24)$$

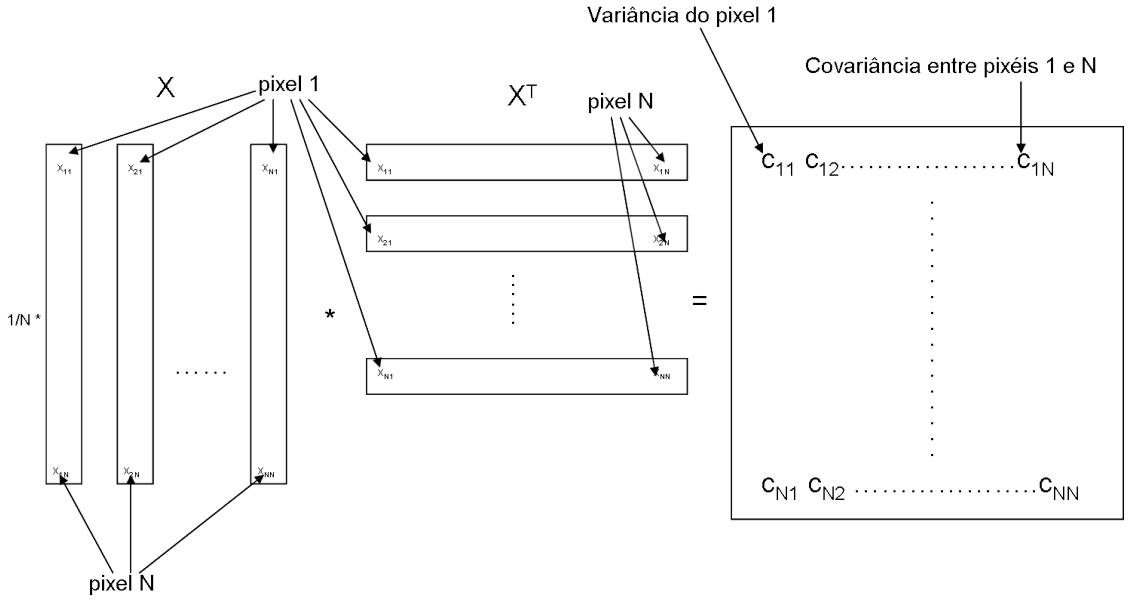


Figura 28 – Matriz covariância

Verifica-se facilmente que os elementos da diagonal da matriz produto $X*X^T$ correspondem à variância de cada pixel, de 1 a N, nas amostras, ou sejas as várias imagens, e que os outros valores correspondem às covariâncias entre pixéis diferentes, por exemplo $C_{100,200}$ será a covariância entre o pixel 100 e o pixel 200, desde que se assuma que as variáveis x_i têm média zero, o que é garantido pelo pré-processamento que descrevemos.

Os componentes U_k são os vectores próprios da matriz C. Como neste processo cada pixel P calculado para cada U_K depende dos pixéis P de todas as imagens da matriz X – ver equação (15) – as imagens que correspondem aos componentes, as chamadas "*eigenfaces*", parecem-se com imagens fantasmagóricas. São imagens de todo o rosto e não podem ser associadas, como tínhamos já visto, a zonas específicas do rosto.

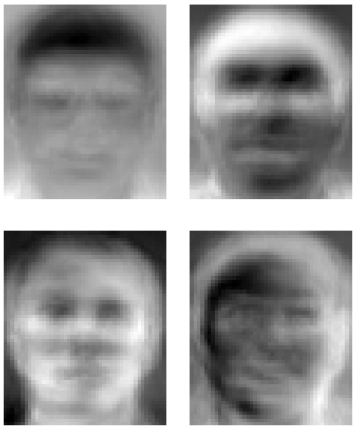


Figura 29 – Eigenfaces da BD ORL

No entanto, se o processo utilizado fosse exactamente o descrito acima, calculando os componentes PCA a partir da matriz covariância C, a dimensão de C seria NxN, o que iria conduzir a N valores próprios e a N vectores próprios. Se pensarmos em imagens de 1000x1000 pixéis, teríamos 10^6, ou seja um milhão de componentes, desaparecendo a vantagem procurada, de passar para um sistema de dimensão muito mais baixa.

Vejamos qual a solução. Tomemos então a matriz S

$$S=X^T*X \qquad (25)$$

Onde X é mais uma vez a matriz das imagens vectorizadas em colunas. S tem dimensão n^2 (n é o numero de imagens, n<<N), em vez de N^2. Calculemos o vectores próprios de S, P_i. Por definição teremos

$$S*P_i=k*P_i \qquad (26)$$

Onde k é uma constante. Ou seja

$$X^T*X*P_i= k*P_i \qquad (27)$$

Ou ainda, multiplicando por X ambos os membros da igualdade

$$(X*X^T)*(X*P_i)= k*X*P_i \qquad (28)$$

A partir deste resultado podemos afirmar que X*Pi são vectores próprios de $X*X^T$. Mas a matriz covariância de X, como vimos antes, é igual a $X*X^T$, logo $X*P_i$ são vectores próprios *de C*.

Portanto o que se faz é calcular S, uma matriz de ordem n^2, em vez de N^2, e calcular os n (P_i) vectores próprios de S. Os primeiros n vectores próprios de C, calculam-se multiplicando os vectores Pi por X, de acordo com a equação (29)

$$Ui=X*P_i \qquad (29)$$

Conseguimos passar de um algoritmo de calculo de ordem N^2 para um algoritmo de ordem n^2 (sendo n tipicamente n << N), que nos permite calcular os n primeiros vectores próprios de C.

Feito este cálculo, obtemos n componentes U_j e, como vimos, cada uma das imagens X_i pode ser calculada a partir desses n componentes U_j, utilizando um conjunto de n coeficientes A_i.

$$X_i^T = A_i * U^T \qquad (30)$$

Os valores A_{ij} representam a projecção de X_i em cada um dos vectores U_j. Esses coeficientes obtêm-se com a equação (31) que calcula o produto interno de X_i por cada um dos vectores U.

$$A_i = X_i^T * U \qquad (31)$$

Se pensarmos em imagens, e em particular em retratos, isto significa que cada vector X_i corresponderia a um retrato. Se a decomposição for feita por exemplo a partir de 200 imagens com 1000x1000 pixéis cada, a matriz A vai ser de 200x200 e cada vector X_i pode ser recuperado a partir dos 200 componentes principais por um conjunto de 200 coeficientes, a linha correspondente a essa imagem na matriz A. Neste cenário o vector X_i, ou seja cada retrato, passa a ser representado por 200 valores em vez de um milhão (número de pixéis), conseguindo-se desta forma a desejada diminuição na dimensionalidade. Em rigor, como veremos, se o número de componentes principais for bastante menor que n, ainda se consegue fazer uma boa recuperação da imagem, em virtude de os valores próprios tenderem normalmente rapidamente para zero. Considerar apenas os componentes mais importantes da decomposição, o que corresponderá a tomar muitos coeficientes de A como nulos, poderá associar-se ao conceito de *sparse coding*.

Naturalmente numa imagem colorida teríamos algo idêntico, para cada uma das três componente RGB.

(Bartlett 1998); (Draper et al. 2003) ; (Hyvärinen and Oja 2000); (Puga 2001) ; (Yambor, Draper, and Beveridge 2000); (Xudong 2006); (Yun 2007);

2.3.4.2 Como reconhecer uma face com PCA

Cada face pode ser representada por um conjunto de coeficientes A_i que representam o seu peso em cada um dos eixos, ou seja para cada um dos componentes do PCA. Portanto uma forma de reconhecer uma face nova será simplesmente comparar os seus coeficientes com os coeficientes das imagens conhecidas que estão guardadas na BD. Em termos vectoriais tal operação corresponderá a determinar a distância ente o vector construído a partir dos coeficientes, obtidos para a nova imagem e os vectores conhecidos da BD. Para tornar possível esta operação será necessário:

1. Construir uma BD de conhecimento de retratos a partir de uma colecção significativa de n retratos
2. Calcular os vectores próprios (*eigenfaces)* para a colecção, tal como se descreveu, e considerar os K vectores próprios mais significativos, em que K <= n
3. Calcular por projecção os coeficientes de cada retrato conhecido num sistema de eixos U, construído a partir dos K vectores próprios obtidos no ponto anterior
4. Perante uma nova imagem, vectorizá-la, eliminar o valor médio e calcular os coeficientes A_i desse vector imagem Y, projectando-o no sistema de eixos U, aplicando a expressão (31), o que resulta na expressão (32)

$$A = (Y - E_Y)^T * U \qquad (32)$$

O vector A dos coeficientes representa a projecção da nova imagem no sistema de eixos U,

correspondente aos K *eigenfaces* mais significativos calculados para a BD, e cada A_i representa o peso relativo de cada um.

5. Procurar na BD a imagem mais próxima da imagem factorizada, utilizando um método de cálculo para o erro, ou para medir a distância entre os dois vectores. Se o erro for < δ a nova imagem corresponderá a esse retrato da BD. Se δ < erro < α o retrato será considerado como um retrato novo que poderá, se pretendermos, ser incluído na BD. Se o erro for > α poder-se-á concluir que a imagem provavelmente não é sequer um retrato.

A definição dos valores de δ e α terá de ser feita durante a configuração e ajuste da aplicação.

A sequência descrita acima, para o caso do PCA, será idêntica para outras técnicas estatísticas descritas nas secções seguintes, a menos de alguns ajustes, pois todas elas transformam a imagem em coeficientes num novo sistema de eixos, escolhidos com base no critério definido por cada uma das técnicas.

(Bartlett 1998); (Draper et al. 2003); (Hyvärinen and Oja 2000); (Puga 2001) ; (Yambor, Draper, and Beveridge 2000); (Xudong 2006); (Yun 2007);

2.3.4.3 Vantagens e limitações do PCA

O PCA é bastante simples e rápido, quando comparado com outros métodos, e permite obter taxas de sucesso de reconhecimento bastante elevadas, mesmo perante alterações como desfoque, ruído, ou alterações do fundo.

O principal objectivo e o critério da transformação feita é no PCA reduzir a dimensionalidade dos dados. A forma como é implementado torna-o muito eficiente em técnicas de reconstrução, com o erro minimizado.

É no entanto muito sensível a outros factores de perturbação tais como alterações de iluminação ou oclusão de parte da imagem. Este último aspecto seria expectável, tendo em atenção o facto de se tratar de uma representação global, onde não há qualquer relação dos componentes com zonas específicas do rosto.
Verifica-se ainda que a qualidade dos resultados obtidos com o PCA depende bastante do cuidado que se tiver no alinhamento das imagens que integram a BD.

(Bartlett 1998); (Bolme et al. 2003); (Draper et al. 2003); (Hyvärinen and Oja 2000); (Yambor, Draper, and Beveridge 2000); (Xudong 2006); (Yun 2007);

2.3.5 Independent Component Analysis – ICA

As técnicas de ICA (Independent Component Analysis) generalizam as técnicas utilizadas como base para o PCA. O ICA tem em conta as dependências de ordem superior nas variáveis X_i, para além das dependências de ordem 2 – as chamadas covariâncias – ou seja, o ICA é um método que necessita de informação associada à distribuição de X, que não está contida na matriz de covariância. O objectivo do ICA é de encontrar componentes independentes que satisfaçam um determinado critério.

Como veremos este método, mais complexo que o PCA, revela-se muito útil quando existem nos dados características não-gaussianas que se revelam importantes para o que pretendemos realizar. A Figura 30 ilustra este facto. Num conjunto de dados como o representado, o 1º eixo

da decomposição PCA seria o eixo vertical, que corresponde à direcção de máxima variância das projecções. Esta solução no entanto esconderia a divisão dos dados em dois clusters, que as funções distribuição evidenciam, e portanto não seria a indicada se essa separação for significativa e se pretender evidenciar. O ICA, pelo contrário, escolheria como primeiro o eixo horizontal, permitindo evidenciar os dois conjuntos independentes de dados.

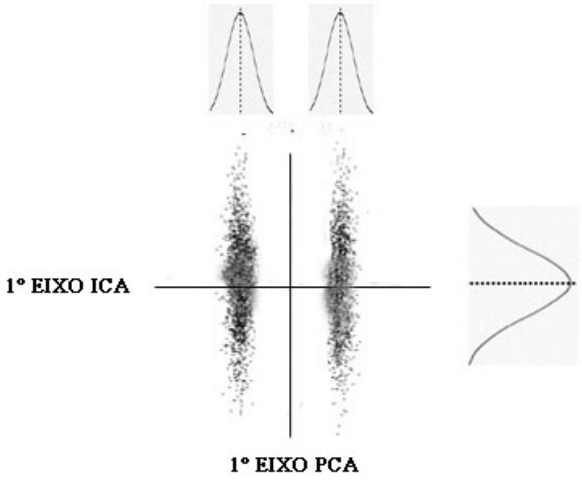


Figura 30 – Exemplo de eixo principal obtido por PCA e ICA

Para a complexidade adicional do ICA fazer sentido, teremos de estar em presença de variáveis não-gaussianas, porque em caso contrário a matriz de covariância conteria toda a informação estatística.

Comecemos por ver o conceito de independência e alguns métodos clássicos, anteriores ao ICA. Veremos depois a relação do ICA com cada um deles.

(Bartlett 1998); (Hyvärinen 1999, 1999); (Hyvärinen and Oja 2000); (Hyvärinen, Oja, and Juha 2001); (Puga 2000)

2.3.5.1 Conceito de independência

Conceito de independência: diz-se que variáveis x_1, x_2, x_n são independentes se a função densidade de probabilidade

$$f(x_1,x_2,....x_n)=f(x_1)*f(x_2)*...f(x_n) \qquad (33)$$

Este conceito é claramente distinto e mais exigente do que o conceito de não correlacionado que apenas impões que

$$E[x_i,x_j]=E[x_i]*E[x_j] \qquad (34)$$

Pode-se demonstrar que os dois conceitos são equivalentes se as variáveis forem gaussianas.

(Hyvärinen 1999, 1999); (Hyvärinen and Oja 2000); (Hyvärinen, Oja, and Juha 2001);

2.3.5.2 Métodos clássicos estatísticos de ordem > 2

Existem três métodos clássicos para tratar informação estatística, utilizando análises de ordem superior a 2, e que antecederam o ICA.

Projecção interessante (*Projection porsuit*) – este método procura identificar o sistema de eixos que conduzem a projecções mais interessantes ou significativas dos dados, de acordo com um determinado critério. Este método é ideal para observar o agrupamento (*clustering*) dos dados. O objectivo de reduzir a dimensionalidade não é tão importante aqui como poder visualizar certas propriedades dos dados, tal como a Figura 30 ilustra. Reduz-se a dimensão, mas assegura-se sobretudo que algumas propriedades interessantes são evidenciadas.

O grande problema teórico destes métodos é definir essa função $f_{interessante}()$ para cada direcção no espaço. Normalmente esta função mede de alguma forma a quantidade não-gaussiana das projecções nessa direcção.

O objectivo vai ser obter a matriz A de coeficientes para calcular U – expressão (35) – de tal forma que a função $f_{interessante}(U)$ seja maximizada.

$$U = X*A \qquad (35)$$

Redução da Redundância – estes métodos procuram precisamente minimizar a redundância dos dados. De acordo com Barlow, que já referimos várias vezes, a redundância parece desempenhar um papel crucial na percepção. Por outro lado Barlow também realçou a importância da independência dos componentes representados, que é essencial, na sua opinião, nomeadamente, para os processos de aprendizagem não supervisionada.

Uma forma de implementar este método consiste em procurar uma solução com *sparse coding*. Neste caso procura-se uma solução em que apenas um número mínimo de componentes seja activado, tal como Barlow acreditava que acontece no cérebro. Se os dados têm propriedades estatísticas que permitem esta aproximação (se os dados são esparsos), este método permite diminuir a redundância.

Desconvolução cega – este método é bastante diferentes dos dois anteriores no sentido em que lida com sinais unidimensionais no tempo em vez de vectores amostra de dados. Neste método, a partir de um sinal convolução x(t), pretende-se obter o sinal fonte s(t), ou seja obter um filtro h(t), tal que

$$s(t) = h(t)*x(t) \qquad (36)$$

A Figura 31 ilustra o processo de desconvolução, ou seja recuperar o sinal original a partir da convolução.

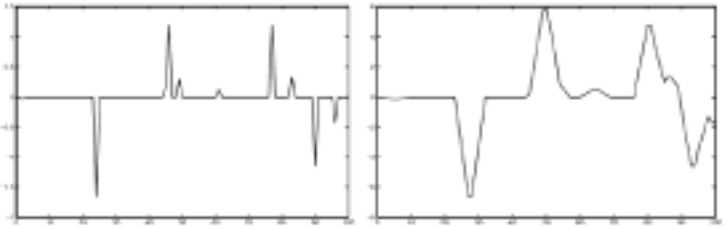


Figura 31 – Recuperar o sinal fonte de um sinal convolucionado

(Hyvärinen 1999, 1999); (Hyvärinen and Oja 2000); (Hyvärinen, Oja, and Juha 2001);

2.3.5.3 Algoritmos ICA e o algoritmo FastICA

Nalguma literatura do ICA encontramos a referência a duas arquitecturas, quase sempre designadas como arquitectura I e II. Na primeira cada imagem é uma amostra e cada pixel uma variável X_i. Na arquitectura II faz-se o contrário: os pixéis 1,2,…,N são as N amostras e as variáveis são o valor de cada pixel em cada imagem. Nós vamos apenas trabalhar com a arquitectura I. Por um lado é a arquitectura ICA mais vezes utilizada e mais popular no reconhecimento de rostos, por outro lado é suposta conduzir a alguma localização física dos atributos nos componentes e interessa-nos contrapor/complementar este aspecto com o carácter global do PCA, nomeadamente no sistema híbrido (PCA+ICA) que pretendemos construir e avaliar. A arquitectura II do ICA, pelo contrário, conduz a imagens nos componentes sem localização, à semelhança do que acontece nos *Eingenfaces* do PCA.

Todos os algoritmos ICA incluem um pré-processamento onde, para além de se centrar os dados em zero, é realizada uma transformação linear para "branquear" (*whitening*) os dados. Esta transformação pretende obter vectores X_i não correlacionados e com variância igual a 1. Ou seja encontrar a partir de X uma matriz X^1 tal que

$$X^{1T} * X^1 = I \qquad (37)$$

Esta transformação é idêntica à realizada na transformação PCA, calculando os vectores próprios da matriz covariância de X. Os componentes ICA vão ser calculados a partir da matriz X^1, tal como se observa na equação (38).

$$X^{1T} = A^1 * U^T \qquad (38)$$

A matriz U é neste caso a matriz dos componentes ICA a calcular, com base num critério a definir. As transformações realizadas simplificam o cálculo e é fácil calcular a matriz de coeficientes pretendida A, a partir de A^1, para reconstruir os dados originais X.

Ao fazer a transformação descrita, a nova matriz de coeficientes A^1 passa a ser ortogonal, reduzindo o número de parâmetros a calcular, pois numa matriz ortogonal teremos, em vez de N^2, apenas Nx(N-1)/2 parâmetros. Uma outra vantagem consiste em poder eventualmente desprezar valores próprios muito pequenos, resultantes da aplicação deste método (*sparse coding*), e assim simplificar ainda mais o cálculo.

Podem ainda existir pré-processamentos que dependem do tipo de dados a analisar, com transformações adequadas à simplificação do método.

Existem diversos métodos ICA, que têm sido propostos como resultado de muitos trabalhos de investigação. Cada método pode diferenciar-se por dois aspectos:

- função objectivo
- algoritmo de convergência

Método ICA= função objectivo + algoritmo

Os dois aspectos são distintos e podemos ter a mesma função objectivo, utilizando algoritmos diferentes, ou eventualmente utilizar o mesmo algoritmo para funções distintas. As propriedades estatísticas dos dados, que o método ICA permite evidenciar, dependem sobretudo da função objectivo. O algoritmo permite sobretudo optimizar e acelerar a convergência do método.

Existem métodos que calculam todos os componentes ICA simultaneamente e que convergem para a solução global, outros que calculam os componentes um a um. Os segundos demonstraram uma maior eficiência. É o caso do algoritmo fastICA, desenvolvido por Aapo Hyvärinen da Universidade

de Tecnologia de Helsínquia, e que utilizaremos no nosso trabalho. O SW é *open source* e o fastICA é considerado como um algoritmo muito eficiente, sendo referido em inúmeros trabalhos que utilizam e testam o ICA para o reconhecimento de rostos. É um algoritmo iterativo que calcula os componentes um a um, encontrando cada direcção U que maximiza o carácter não-gaussiano das projecções de X ($X*U^T$). Utiliza uma aproximação da função de negentropia para medir o carácter não-gaussiano da projecção dos dados.

O ICA não nos diz como ordenar os componentes ICA, ao contrário do que acontecia no PCA, onde a ordenação era natural pelos valores próprios. Um método para fazer essa ordenação consiste em utilizar a norma das colunas da matriz de mistura A, pois as normas representam a contribuição de cada componente na reconstrução dos vectores X_i. É um método de ordenação idêntico ao utilizado no PCA. Outro método consiste em tentar medir de alguma forma o carácter não-gaussiano de cada coluna, para fazer essa ordenação.

(Bartlett 1998); (Draper et al. 2003); (Hyvärinen 1999, 1999); (Hyvärinen and Oja 2000); (Hyvärinen, Oja, and Juha 2001);

2.3.5.4 Vantagens e limitações do ICA

O ICA é um método diferente do PCA e tem objectivos diferentes. Também permite uma redução da dimensionalidade dos dados, mas o principal objectivo é o de escolher as direcções, que permitem maximizar uma determinada função, definida com base num critério que irá depender doa aspectos que se pretendem realçar com esta transformação.

É mais complexo e alguma investigação indicia que pode ser menos eficaz do que o PCA, em termos de reconstrução das imagens com erro mínimo, como constatámos no nosso trabalho. No entanto a literatura é contraditória nos resultados, quando compara os dois métodos quanto à eficácia no reconhecimento de retratos. Uma razão para esta aparente contradição é termos muitos parâmetros a considerar, quando se faz essa comparação, outra é o facto da eficiência e eficácia do ICA dependerem da função utilizada, da forma de calcular o erro e do algoritmo utilizado. Uma das razões para um grande interesse no ICA resulta de este método (arquitectura I), como referimos, permitir obter componentes que parecem orientadas a zonas da imagem e por isso poder ser útil em processos em que uma imagem é considerada como composta por várias unidades, por exemplo, olhos, nariz, boca, etc. Tal como se reclama que acontece com o NMF, como veremos.

O ICA tem inúmeras aplicações:

Separação de sinais: suponhamos que temos vários sinais áudio misturados, o exemplo clássico será termos um conjunto de microfones numa sala que gravam o somatório das vozes de todos os presentes. O sinal captado em cada microfone **i**, $x_i(t)$, será a soma de n componentes, relativos a cada um dos sinais originais **s**, tal como se apresenta na equação (39).

$$x_i(t) = a_{i1}s_1 + a_{i2}s_2 + \ldots a_{in}s_n \qquad (39)$$

Recuperar os sinais seria fácil, se se conhecesse os valores dos coeficientes a_{ij}. Não os conhecendo, como normalmente acontece, é uma tarefa complexa. Uma aplicação suportada em ICA é capaz no entanto de separar esses sinais e recuperar cada uma das vozes, desde que se assuma que os sinais são independentes.

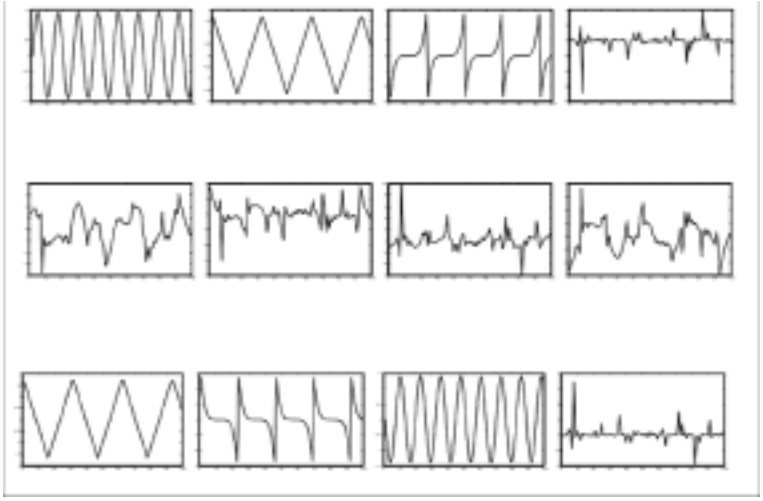


Figura 32 – Separação de sinais misturados

Outro exemplo parecido poderia ser pretender separar um sinal, por exemplo uma voz, de ruído, que a torna imperceptível.

Estas aplicações têm tido um grande desenvolvimento na área da medicina, por exemplo na separação de um conjunto de sinais na electroencefalografia ou na magnetoencefalografia.

Extracção de propriedades: neste caso os componentes ICA representam propriedades que pretendemos observar. Existem trabalhos, que procuram extrair as propriedades de imagens naturais, utilizando o conceito de "sparse coding", que referimos. Esta aproximação tem como um dos objectivos diminuir a redundância dos dados. Alguns resultados parecem validar a semelhança com resultados observados no córtice visual primário de um macaco.

Outras aplicações: as características do ICA e em particular a sua vocação para encontrar propriedades específicas nos dados, procurando direcções que correspondam a essas propriedades, fazem prever poderem existir inúmeras áreas onde este método pode vir a ser utilizado com sucesso.

(Bartlett 1998); (Bolme et al. 2003); (Draper et al. 2003); (Hyvärinen 1999, 1999); (Hyvärinen and Oja 2000); (Hyvärinen, Oja, and Juha 2001); (Yambor, Draper, and Beveridge 2000); (Wendy S. Yambor 2002);

2.3.6 Non-Negative Matrix Factorization – NMF

Os métodos anteriores – PCA e ICA – permitem uma redução da dimensionalidade das imagens. Mas ambos conseguem grande eficiência nessa redução destruindo em parte a estrutura física com significado que suporta as imagens, o que pode dificultar a análise e transformações adicionais. Esta dificuldade, na opinião dos proponentes do NMF, deriva sobretudo do facto de serem métodos numéricos, que resultam no aparecimento de coeficientes negativos, sem qualquer contextualização física.

O NMF foi proposto em 1999 por Lee e Seung_como um método novo de factorização de dados, tal como o PCA ou o ICA, mas com restrições de não negatividade. De acordo com os proponentes dos NMF, o método permite, tal como o PCA ou o ICA, a redução na dimensão dos dados, mas sem destruir na transformação a estrutura física por detrás das imagens. Portanto, de acordo com os proponentes deste método, uma das vantagens poderia ser encontrar componentes que correspondem a características físicas, bem localizadas na imagem, tal como se observa na Figura

33, onde, em vez das imagens globais fantasmagóricas do PCA, teríamos para cada componente NMF uma zona específica das imagens de que partimos.

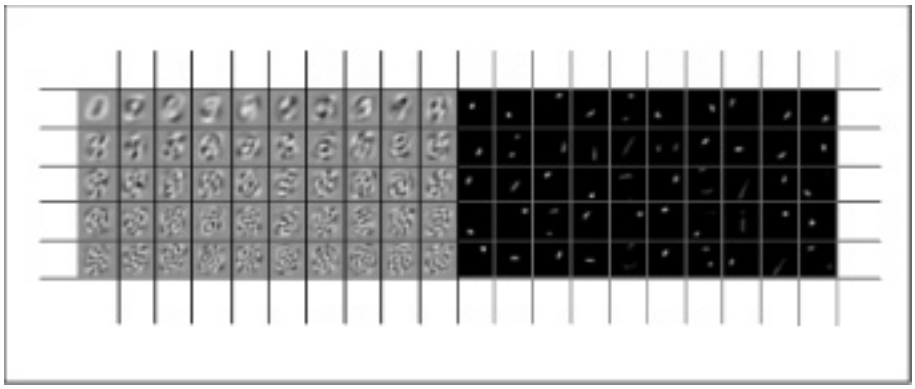


Figura 33 – Imagens de componentes obtidas com PCA e com NMF

Vejamos como funciona o método NMF. Tal como no PCA ou ICA, as imagens são vectorizadas para construir uma matriz X de n imagens, com dimensão N x n (n << N), onde n é o número de imagens e N o número de pixéis por imagem. Para fazer a factorização desejada, aproxima-se a matriz X ao produto das duas matrizes, W e H, onde W é uma matriz de r componentes NMF e H uma matriz de coeficientes, idêntica à matriz dos coeficientes do PCA ou do ICA. No método impõe-se a não negatividade de todos os elementos.

A matriz X é portanto aproximada ao produto de duas matrizes.

$$X = (W*H) \qquad (40)$$

Se r for o número de componentes NMF, a matriz W tem uma dimensão Nxr e a matriz H uma dimensão rxn. Normalmente o valor de r é tão pequeno quanto possível, tipicamente r <= n, em que n é o número de imagens. Ao contrário do PCA ou do ICA, onde os valores próprios podem justificar a dimensão r a adoptar ao transformar, no caso no NMF a escolha de um valor adequado para r é ainda um problema mal resolvido.

O segredo dos algoritmos NMF está nos valores iniciais das matrizes e no algoritmo iterativo que permite aproximar o produto das duas matrizes à matriz de dados, debaixo da restrição de não-negatividade. Muitos trabalhos de investigação recentes têm proposto novos e pretensamente mais eficientes algoritmos, para calcular as matrizes W e H.

O objectivo do algoritmo será tentar minimizar a função *f(W,H),* ou seja a soma das diferenças quadráticas entre X e a aproximação NMF.

$$f(W,H) = \sum_{i=1}^{n}\sum_{j=1}^{n}(X_{ij} - (W*H)_j)^2 \qquad (41)$$

Alguns autores propuseram uma função de medição do erro da aproximação, para orientar a convergência, diferente da função das diferenças quadráticas: a função divergência de Kullback-Leiblerentre para X e W*H.

$$\min_{W,H} \sum_{i=1}^{n} \sum_{j=1}^{m} (X_j \log {X_j}/{(W*H)_j} - X_j + (W*H)_j) \qquad (42)$$

Mas a solução NMF não é única. Se (W,H) é uma solução, então é fácil concluir, a partir da equação (40), que por exemplo ($W*D, D^{-1}*H$) também é uma solução.

Existem vários algoritmos que utilizam métodos diferentes para fazer a factorização NMF.

O método mais conhecido e referido em todos os trabalhos que utilizam NMF é o método de actualização multiplicativo, proposto por Daniel Lee e Sebastian Seung, os primeiros autores a propor o NMF. Neste método as matrizes W e H são inicializadas aleatoriamente com valores positivos e vão sendo multiplicados por factores, calculados de forma a se convergir para a solução. Os factores de multiplicação são apresentados nas equações (43 e 44)

$H_{k+1} = H_k*(W_k'*X)/((W_k'*W_k)*H_k+eps);$ (43)[4]
$W_{k+1} = W_k*(H_k*X')'/(W_k*(H_k*H_k')+eps);$ (44)
$X_R = W*H;$ (45)

erro (distância euclidiana)= $\sum_{i,j}(X_{ij} - X_{Rj})^2$ (46)

O processo é iterativo e pára para um determinado limiar do erro (que pode ser calculado utilizando diferentes definições, por exemplo através da distância euclidiana - ver equação (46) – ou para um número máximo de iterações. A convergência deste método, ainda que reclamada pelos autores, é ainda objecto de alguma polémica.

Um outro método, utilizado por alguns algoritmos, é o método de mínimo quadrados alternado (Alternating Least Squares). Neste caso cada iteração de uma das matrizes (H,W) é calculada da última iteração da outra matriz. Ou seja, alternadamente, uma matriz é optimizada, mantendo a outra fixa. Considerando as duas matrizes W e H o problema de minimizar a função erro é não linear e complexo, no entanto se fixarmos uma das matrizes o problema transforma-se num problema trivial de minimização.

Neste método a matriz W é inicializada aleatoriamente. A partir de W_1 calcula-se H_1, a partir de H_1 calcula-se W_2 e assim sucessivamente

$\min|X - W_1H|$ - fixando W_1, permite calcular H_1
$\min|X - WH_1|$ - calculado H_1, permite calcular W_2

O processo de cálculo de H_k e W_k é iterativo, até o nº de iterações ou o erro ultrapassarem o limiar imposto.

O mínimo para $|X - WH|$ em cada iteração é calculado decompondo o problema em n problemas, ou seja, encontrar mínimos para cada coluna j da matriz erro

$\min|X_j - WH_j|$

Minimizar cada vector coluna corresponde a minimizar a soma dos quadrados dos seus termos (donde provém o nome destes algoritmos - least squares). É um problema conhecido para o qual existem algoritmos, nomeadamente no matlab (lsqnonneg).

Nos últimos anos, foram no entanto propostos algoritmos diferentes, que utilizam como referência a teoria da optimização para problemas com restrições, a que no fundo se pode reduzir o problema da factorização de matrizes NMF. Nestes problemas pretende-se minimizar uma função f(x), colocando restrições às variáveis $x_i \in R^n$, tal como se define em (47)

$$\min f(x) \rightarrow v_i < x_i < u_i \qquad (47)$$

No fundo, é o problema que temos no NMF, onde f(x) vai ser uma função que mede o erro da

4 Nestas expressões X', W' e H' identificam as matrizes transpostas

aproximação e onde todos os termos das matrizes W e H têm de ser > 0. Ora uma das soluções existentes para este tipo de problemas recorre aos gradientes projectados. Estes algoritmos NMF utilizam portanto os valores dos gradientes projectados para iterativamente calcular W_{k+1},H_{k+1} a partir de W_k,H_k. Estes algoritmos são apresentados pelos autores como mais rápidos e mais eficientes do que aqueles que recorrem ao método multiplicativo.

Iremos neste trabalho experimentar vários algoritmos NMF, cobrindo o espectro dos métodos propostos, utilizando uma toolbox disponibilizada pela Universidade Técnica da Dinamarca. Esta toolbox integra:

- o primeiro algoritmo proposto para a factorização NMF, de actualização multiplicativo, proposto por Daniel Lee e Sebastian Seung
- o algoritmo de mínimos quadrados alternados
- o algoritmo de gradientes projectados, proposto por Chih-Jen Lin

No caso do algoritmo de gradientes projectados vamos ainda testar directamente o algoritmo original proposto por Chih-Jen Lin

(Lee and Seung 2001); (Zafeiriou, Tefas, and Pitas 2005); (Zafeiriou et al. 2005); (Chih-Jen 2007); (Berry et al. 2007); (Yun 2007)

2.3.7 Linear Discriminant Analysis - LDA

O método LDA (Linear Discriminant Analysis) também conhecido como FLD (Fisher Linear Discriminant) utiliza o conceito de classe, para agrupar retratos e dessa forma conseguir melhorar a performance do reconhecimento. É um método que se distingue dos três anteriores por corresponder ao que se chama um método supervisionado de aprendizagem *("supervised learning"*), pois obriga a uma classificação prévia das imagens, para as poder agrupar em classes, para por exemplo agrupar todos os retratos de uma mesma pessoa numa mesma classe. Esta aprendizagem supervisionada e portanto não automática permite separar no espaço N-dimensional todos os pontos de cada uma dessas classes por um hiperplano e esse hiperplano permite classificar no futuro, já sem supervisão, novas imagens, determinando a(s) classe(s) a que pertencem. Este método permite assim classificar novos rostos, por exemplo como rostos de mulher ou de homem, ou como rostos de latinos ou de nórdicos, se tivermos definidos os hiperplanos que separam essas classes.

Apesar de não termos tido tempo, na realização do nosso trabalho, para introduzir e experimentar este método, descrevemo-lo sumariamente aqui porque acreditamos ser um método com interesse para futuros desenvolvimentos, pois, ao introduzir o conceito de classe, permite jogar com o conceito de identidade associado a uma classe de retratos.

A grande aplicação deste método tem sido para melhorar a eficiência e a eficácia no reconhecimento em situações onde as alterações de iluminação e as expressões dos retratados podem ser problemáticas para o PCA ou ICA.

No LDA pretende-se construir um sistema que discrimina muito bem entre classes e não entre imagens individualizadas, ao contrário por exemplo do PCA que pretende maximizar a discriminação entre cada imagem. O algoritmo procura uma base de vectores que maximize a relação da variação entre imagens de classes distintas. Ou seja, pretendem-se encontrar vectores que correspondam às direcções predominantes nas diferenças entre classes e que sejam se possível quase ortogonais às direcções que correspondem às variâncias máximas dentro de cada classe e assim tornar tanto quanto possível o sistema imune às variações internas na classe, por exemplo independentes da iluminação. Verifica-se que muitas vezes as variações devidas a alterações na iluminação, como se observa na Figura 34, são mais significativas do que as variações resultantes de termos duas pessoas diferentes e este facto é muito mal resolvido pelo PCA ou pelo ICA.

Figura 34 – Dois retratos da mesma pessoa com iluminações diferentes

A diferença entre LDA e PCA é ilustrada pela Figura 35, onde, a partir das observações registadas para as duas classes, classe 1 e 2, o eixo principal num sistema baseado em PCA e num sistema baseado LDA, seriam completamente distintos.

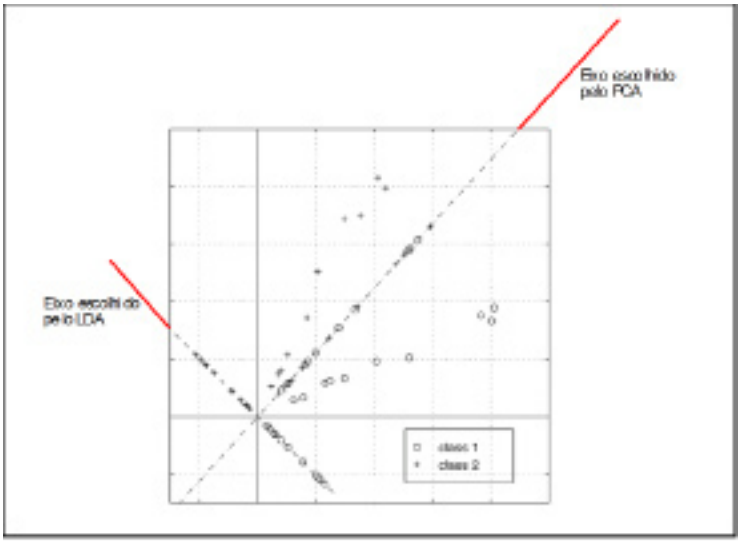


Figura 35 – Eixos encontrados para duas classes em PCA e LDA

O PCA tenderia a maximizar a discriminação entre pontos individuais, enquanto o LDA tentaria maximizar a discriminação entre classes e minimizar a discriminação entre pontos da mesma classe. Por isso em LDA o eixo encontrado seria, neste exemplo, ortogonal ao eixo PCA. Seria assim fácil neste sistema LDA saber a que classe cada ponto teria mais probabilidade de pertencer. Bastaria verificar se U_1, o coeficiente de projecção no eixo escolhido, era positivo ou negativo.

É como se o algoritmo que suporta o LDA tendesse em parte a valorizar as partes do rosto que se mantêm nas várias imagens de uma classe e que porventura são mais importantes na identificação nestas condições e a desvalorizar outras partes do rosto ou outros aspectos que variam muito de imagem para imagem na mesma classe.

A vantagem do LDA é portanto a imunidade a factores de variação como a iluminação ou a expressões diferentes. Permite ainda, independentemente desses factores, elevada eficácia no reconhecimento de imagens.

O LDA é no entanto bastante mais pobre que o PCA na reconstrução de uma imagem pois a importância de cada imagem em parte é perdida.

A complexidade do LDA obriga, na maioria das abordagens deste método, a utilizar o PCA como pré-processamento, para conseguir reduzir a dimensionalidade dos dados, antes de utilizar os algoritmos LDA.

Alguns autores propuseram entretanto a utilização de um método em que o NMF é misturado com o LDA: durante o algoritmo NMF, que é um processo iterativo, a informação sobre as classes criadas é tida em conta, ao escolher a iteração seguinte, procurando optimizar a discriminação interna e externa ao mesmo tempo que impõe a positividade das matrizes.

(Zafeiriou et al. 2005; Zafeiriou, Tefas, and Pitas 2005); (Xudong 2006); (Yun 2007)

– Pois é disso que me queixo – disse Humpty Dumpty – A tua cara é a mesma de toda gente, então... – (marcando os lugares no ar com o polegar) – ... nariz no meio, boca por baixo. É sempre a mesma. Agora, se tivesses os dois olhos ao lado do nariz... ou a boca por cima... já seria de alguma ajuda.

Lewis Carrol

A CABEÇA ETRUSCA, ALFABÉTICA

a) **UMA ESPÉCIE DE LUMINOSIDADE**

Captado pela luz um rosto é
uma espécie de luminosidade
mais intensa, a inquietude
reclusa de um olhar sem
legendas ou na penumbra
mais geográfica do silêncio –
um homem está parado no
centro desta imobilidade,
escuta o vento, as narinas
abrem-se, a saliva desliza pela
língua e pára na garganta
sêca, os lábios gretam o
mundo, o seu olhar que nos
observa resiste ao destino do
tempo – entre ruínas e
sujidade o tédio avança a sua
usura e resiste, um osso de
luz deflagra na paisagem os
seus escombros, deciframos
inéditos instantes pela
melodia rugosa das palavras
que um rosto escuta, pelo
bolor que estanca o sangue
nas carótidas e nos joelhos.

b) **O DELÍRIO NEGRO DAS PALAVRAS MAIS EXPLOSIVAS**

É um veneno o pensamento refreando a gangrena
a paisagem bruta ou a sonda de uma cegueira
espécie desgastada de lume que se escoa
na vastidão das películas no obscuro
equilíbrio que se desdobra das texturas
até ao êmbolo de um olhar ou agulha
que estanca o sangue cozida a cicatriz
os minuciosos túneis que fazem gelar
o delírio negro das palavras mais explosivas
aluem e na agonia a loucura redime
a noite e freme na palavra de deus
um caminho antigo e oblíquas
as sombras fervem na íntima turbulência
no pavor inóspito e na fértil luz
que a rigorosa geometria da dança
desenha no vulcânico ritmo de uns lábios
insonoros – falemos então da morte
agora que a beleza brilha.

Jorge Velhote

3 Metodologia

Esta secção descreve a metodologia adoptada para validar e comparar os resultados obtidos ao decompor e reconstruir retratos, com cada uma das técnicas estatísticas avaliadas no nosso trabalho.

Descreve-se a BD de retratos utilizada no projecto e define-se o conceito de erro de reconstrução utilizado ao longo de todo o trabalho.

A avaliação e a comparação das diversas técnicas foi realizada em quatro fases:

A. Escreveram-se pequenos scipts Matlab, suportados em algoritmos públicos, que permitiram validar cada uma das técnicas estatísticas avaliadas, calculando os respectivos componentes a partir da BD de retratos. Utilizámos um retrato da BD e também um retrato externo do grupo de testes, para medir e comparar os resultados obtidos na reconstrução desses retratos, em particular o erro de reconstrução. Conseguimos assim validar cada uma das técnicas e cada algoritmo. Validámos de forma idêntica sistemas híbridos que associam PCA, ICA e NMF e um algoritmo que desenvolvemos para estes sistemas não ortogonais.

B. Depois escreveram-se scripts Matlab para avaliar o comportamento médio de cada sistema, permitindo calcular valores médios globais – nomeadamente para o erro de reconstrução – para o conjunto de todos os retratos. Esse comportamento médio foi avaliado para os 400 retratos da BD e para os 39 retratos externos de teste, permitindo dessa forma comparar os comportamentos médios para cada técnica estatística experimentada e uma vez mais para sistemas híbridos que associam componentes PCA, ICA e NMF.

C. Os retratos foram posteriormente tratados, para obter uma BD normalizada, com todos retratos coincidentes na localização dos dois olhos – critério adoptado – para comparar os resultados obtidos com as duas versões da BD – a BD inicial e a BD assim normalizada.

D. Os retratos reconstruídos foram utilizados como suporte para um inquérito, para determinar a capacidade de reconhecimento humano de um retrato reconstruído, em função do erro na reconstrução.

3.1 Construção da BD de retratos

Foi construída uma BD de retratos de faces para o projecto. Os retratos são todos frontais, à mesma distância e com as mesmas condições de iluminação. Para realizar esses retratos organizaram-se várias sessões fotográficas em diversos locais, onde se pediu a colaboração de alunos, professores, investigadores e funcionários. A Tabela 2 resume estas sessões.

Local da Sessão	Nº pessoas fotografadas
FEUP	106
INESC	85
FEUP	104
FEUP	61
ISEP	79
OUTRO	4
Total	439

Tabela 2 – Sessões fotográficas

Os retratos foram tratados em Photoshop. Foram limpos e foi feito um pequeno ajuste de contraste e de níveis.

Foi construída uma BD de conhecimento com 400 retratos a preto e banco e uma BD de teste com 39 retratos também a preto e branco, todos com 200x200 pixéis. Esta resolução foi condicionada pela memória disponível e pela memória necessária ao processamento, que envolve o cálculo e manipulação de matrizes enormes. Para o projecto artístico adoptaram-se imagens na BD com uma resolução de 600x600 pixéis, para conseguir a qualidade pretendida na impressão, tendo sido por isso necessário recorrer a uma máquina com a memória que esse cálculo exigia.

*Nota: como não dispúnhamos ainda da nossa BD de retratos, os primeiros resultados para PCA e ICA foram obtidos ainda com uma BD pública obtida da net chamada **faces94**. É uma BD de 110 imagens com 200x180 pixéis.*

http://cswww.essex.ac.uk/mv/allfaces/faces94.html

Os resultados foram depois confirmados e aprofundados com a nossa BD.

3.2 Conceito de erro de reconstrução

Ao reconstruirmos um retrato a partir dos componentes PCA, ICA e NMF, existe um conceito fundamental que é o conceito de erro no processo de reconstrução. No nosso trabalho vamos utilizar dois conceitos de erro: o erro geométrico que pode ser medido utilizando múltiplas definições matemáticas[5], embora todas elas de alguma forma meçam a diferença entre os dois vectores, a

5 Encontrámos na literatura 19 definições para o erro:
(1) Manhattan distance (L1 metric, city block distance)
(2) Euclidean distance (L2 metric)
(3) Chebychev distance (L-∞ norm)
(4) Mahalanobis distance
(5) Lance distance
(6) Statistical distance
(7) Divergence
(8) Kullback-Leibler distance (Relative Entropy)
(9) Symmetrized divergence
(10) Symmetrized Kullback-Leibler distance
(11) Earth mover's distance (EMD)
(12) Mahalanobis angle distance
(13) Chi square distance
(14) Exponential similarity coeficient-based distance
(15) Non-parametric similarity coeficient-based distance
(16) Cosine distance
(17) Correlation coefficient-based distance (CCBD)
(18) Non-negative vector similarity coeficient-based (NVSC1) distance 1
(19) Non-negative vector similarity coeficient-based (NVSC2) distance 2

imagem original X e a imagem reconstruída X_R, dois vectores num espaço N-dimensional, e o erro fotográfico que é um erro subjectivo, relacionado com a qualidade da reconstrução fotográfica dos diferentes processos.

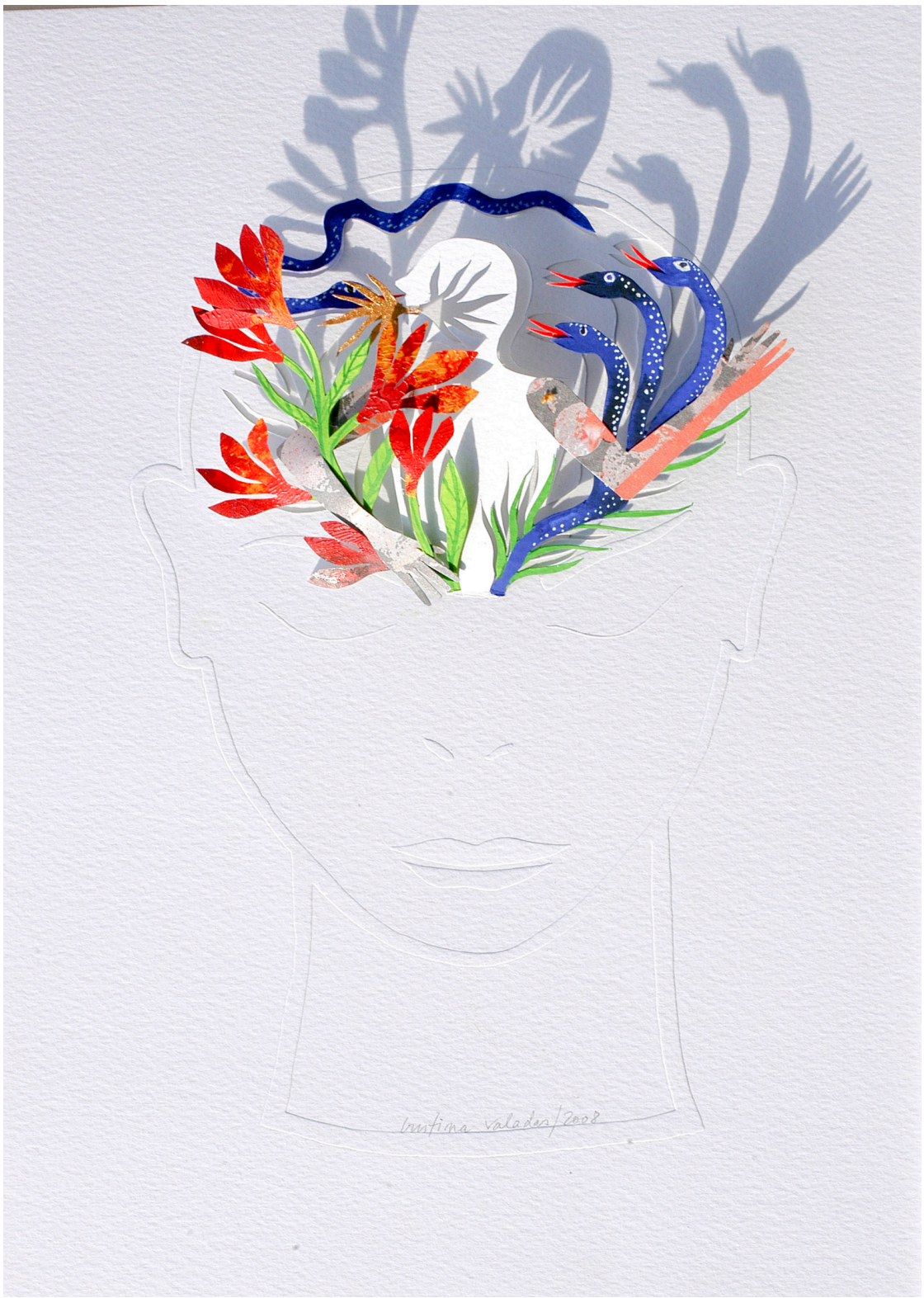
No nosso trabalho o erro geométrico no processo de reconstrução vai ser sempre calculado com base no erro quadrático, a partir do vector diferença entre os dois vectores, tal como se mostra na equação (48).

$$Erro = \sum_i (X_i - X_{Ri})^2 \qquad (48)$$

Este valor, que corresponde ao módulo do vector erro, mede de alguma forma a energia relativa contida nesse erro, ou se preferirmos a relação ruído/sinal, pois os vectores imagem X_i vão ser sempre todos normalizados com comprimento unitário. No trabalho, por facilidade, decidimos utilizar sempre para este erro uma escala de 0 a 100, onde 100 corresponderia a termos $|erro|=|X_i|=1$.

O conceito de erro fotográfico, associado à reconstrução de um retrato, apesar de subjectivo, também vai ser objecto de avaliação no nosso trabalho. Iremos nomeadamente estabelecer uma relação entre os dois conceitos: erro geométrico e erro fotográfico ou qualidade fotográfica.

(Bartlett 1998); (Yun 2007);



Um retrato! O que pode ser mais simples e mais complexo, mais óbvio e mais profundo?

Charles Baudelaire

R. não poderia em rigor afirmar que fora maltratado. Obrigaram-no a subir as escadas e a entrar naquela pequena sala escura, quase sem pronunciarem palavra. Disseram “Entre!” e ele entrou. Durante muito tempo não conseguiu ver nada para além daquele foco de luz que iluminava uma cadeira à frente de uma mesa. Na mesa estava pousado um livro grosso. Ouviu outra voz, “Sente-se!”. Avançou e sentou-se na cadeira. Só algum tempo depois conseguiu perceber que havia um corpo sentado do outro lado da mesa, quando ouviu de novo a mesma voz a ordenar-lhe em tom neutro “Diga-nos onde está!“. Como R. não reagisse a voz insistiu “ Abra o livro e procure o seu retrato!”. R. não percebia o que pretendiam dele mas fez o que lhe mandaram. Ergueu os braços e folheou o grosso livro pousado na mesa. Devagar, página a página, sem reconhecer ninguém. “Então?”. R. confessou que não identificava nenhum retrato do livro. “É melhor colaborar, pois não sai daqui sem nos dizer quem é!”. R. estava cansado. Tinha dormido pouco. Continuou a folhear o livro enorme, página a página, sem conseguir reconhecer um retrato. Quando se adaptou à luz da sala distinguiu melhor o vulto que o interrogava, que de vez em quando movia as mãos, e um outro corpo que parecia estar de pé, por detrás. As horas passavam e a situação não evoluía. De vez em quando ouvia a mesma voz do outro lado da mesa “Então?” e R. nada. A certa altura o vulto sentado levantou-se e ouviu o som de passos a afastarem-se da mesa. Um ruído do que lhe pareceu ser uma porta a fechar, mais distante, levou-o a pensar que ele saíra da sala. O segundo vulto aproximou-se então de R. e com uma voz velada disse-lhe baixinho “Experimente folhear as folhas rapidamente, como fazia em pequeno na escola, com os cadernos de bandas desenhadas rudimentares”. R. a princípio não compreendeu o que ele pretendia dizer. “Lembra-se do pássaro desenhado no bloco de papel que batia as asas?”. R. lembrou-se, ergueu as mãos para o livro, fez rodar as páginas à velocidade certa e começou a ver retratos conhecidos: o avô e a avó, a mãe a tratar das flores, o pai a fumar cachimbo, R. menino, R. na escola. R. viu uma lágrima a correr-lhe no rosto, viu uma sala escura onde o interrogavam, R. a parar de folhear um livro enorme, pois não aguentava mais. O vulto que saíra regressou ao seu lugar e perguntou como se nem sequer tivesse saído dali “Então? Estamos mais dispostos a colaborar?”. R. não respondeu. Ergueu as mãos para o livro, folheou-o e viu R. a sorrir feliz em frente ao mar.

Reborn Bramble [tradução de Renato Roque]

4 Reconstrução de retratos com a BD inicial

Esta secção vai apresentar e discutir os resultados obtidos com PCA, ICA, NMF e com sistemas híbridos, a partir da nossa BD de retratos inicial, por normalizar. Na secção 5 apresentamos e comparamos os resultados obtidos com a BD normalizada.

4.1 A BD de retratos

Figura 36 – BD com 400 retratos com 200x200 px

Dos 439 retratos que realizámos seleccionámos 400 para a nossa BD de retratos. Os 39 retratos restantes serviram para criar um conjunto de retratos fora da BD, para testes.

As pessoas fotografadas na FEUP, INESC e ISEP ofereceram-se voluntariamente para participar neste projecto. Trata-se de um universo com indivíduos dos dois sexos, ainda que com predominância do sexo masculino – 300 homens e 139 mulheres – valores que reflectem uma ainda maior percentagem de homens no universo da Engenharia, com idades entre os 18 e os 65 anos.

Utilizámos retratos a preto e branco com uma resolução de 200x200 pixéis.

4.2 Resultados com o PCA

Desenvolveram-se um conjunto de scripts em MATLAB, utilizando o algoritmo de PCA que o MATLAB disponibiliza, para fazer a decomposição da matriz de vectores de imagens em componentes PCA.

Apresentamos os principais resultados obtidos para o PCA:

- Os 400 componentes PCA
- O erro na reconstrução de uma imagem
- O processo de reconstrução com K componentes, com K=1:400

Os componentes PCA foram calculados com um pequeno script de Matlab que lê todos os retratos da BD a partir de uma directoria de imagens, os transforma em vectores, lhes retira o valor médio e os normaliza com comprimento unitário, para construir a matriz X de retratos. Depois calcula as matrizes A, E e U, onde cada vector U_i é um vector próprio da matriz covariância de X e representa uma componente PCA, onde E é o vector dos valores próprios e onde a matriz A é a matriz de mistura com os coeficientes PCA. Neste cálculo o script utiliza o algoritmo de PCA do MATLAB. Depois, para cada componente PCA, calcula a matriz imagem associada e guarda-as numa directoria de componentes em formato tiff ou jpg. As matrizes A, E e U são guardadas num ficheiro de dados Matlab.

Os 400 componentes PCA, que calculámos com a nossa BD – ver Figura 37 – utilizando o script Matlab mencionado confirmaram a ideia de corresponderem a imagens fantasmagóricas de faces humanas, onde não é visível qualquer predominância de nenhuma zona da face.

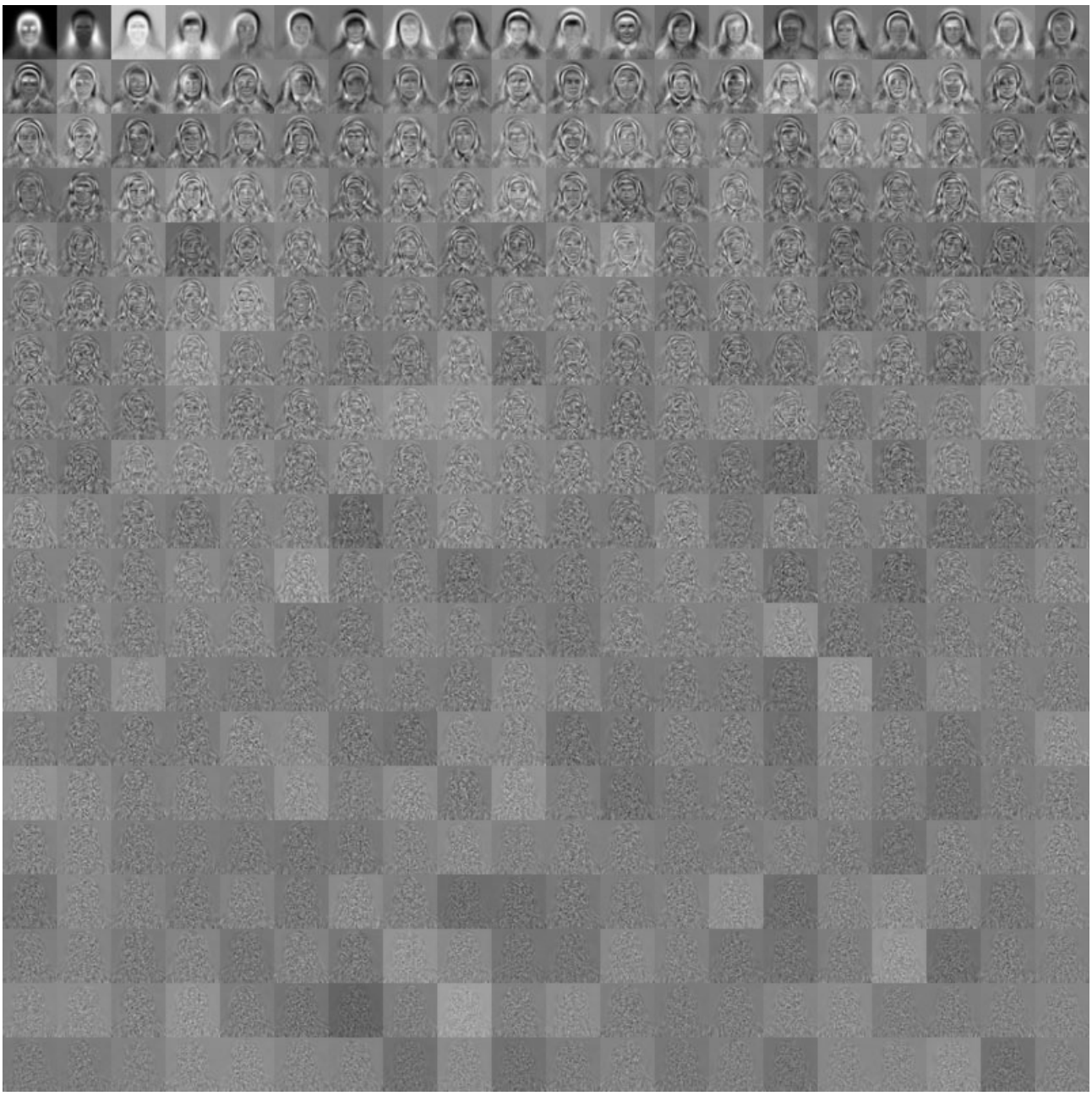


Figura 37 – Os 400 componentes PCA com 200x200 px

Utilizando um outro script Matlab calculámos depois o erro para um retrato reconstruído pelo método PCA, utilizando 1:400 componentes. O script lê o ficheiro com o retrato a testar, vectoriza-o e normaliza-o, retirando a média e dando-lhe comprimento unitário. Obtemos um vector X para esse retrato. Depois, utilizando os coeficientes PCA, o script faz uma reconstrução do retrato, calculando um vector linha X_R por um processo iterativo, utilizando de cada vez mais um componente de 1:400 componentes. Para isso é utilizado um vector de coeficientes A_K, obtido a partir do vector dos coeficientes PCA para esse retrato A_i – A_i corresponde à linha de ordem i da matriz A, a matriz dos coeficientes calculada pelo PCA, se o retrato for o retrato ordem i da BD. A_K é facilmente obtido, em cada iteração K, a partir de A_i, fazendo zero todos os coeficientes de ordem> K, garantindo assim na equação (49) que os componentes de ordem> K não são utilizados na reconstrução.

$$X_R = A_K * U^T \qquad (49)$$

É calculado o erro da reconstrução comparando X_R com X em cada iteração, utilizando a equação

(48); cada vector retrato X_R assim reconstruído é recalculado como uma matriz imagem de 200x200 pixéis e guardado numa directoria de reconstruções com o formato tiff ou jpg.

De realçar a evolução do erro que apresentamos na Figura 38. O erro converge muito rapidamente para zero. O valor final do erro por exemplo para o retrato da BD que testámos é de 4.27 e-024, ou seja, zero. Observa-se que com pouco mais de 20 componentes o erro de reconstrução é inferior a 10 em 100, o que justifica sermos capazes de identificar um rosto familiar com pouco mais de 20 componentes.

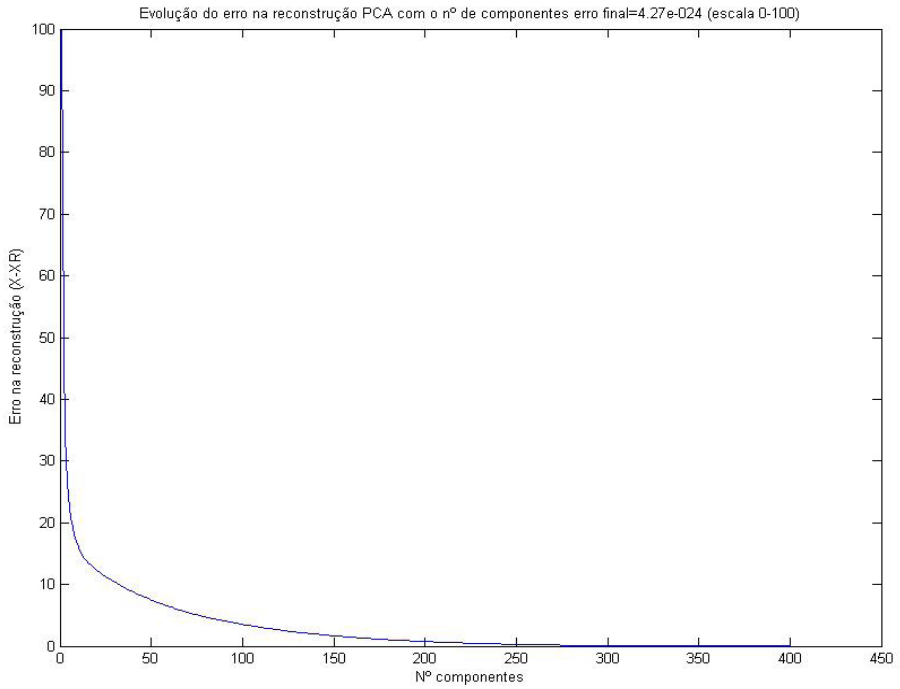


Figura 38 – Erro da reconstrução com K componentes (K=1:400)

Confirmámos portanto a capacidade do PCA para minimizar a dimensionalidade dos dados – neste caso de 40000 para 400 – e para reconstruir as imagens com um erro mínimo. As imagens recuperadas a partir do PCA, com todos os componentes, não se distinguem dos originais e com pouco mais de 20 componentes, como se observa na Figura 38, uma imagem da BD, construída somando as projecções dessa imagem nos eixos PCA, é já relativamente próxima da imagem original.

A Figura 39 e a Figura 40 mostram a evolução de um retrato da BD, reconstruído com K componentes PCA, com K de 1:400.

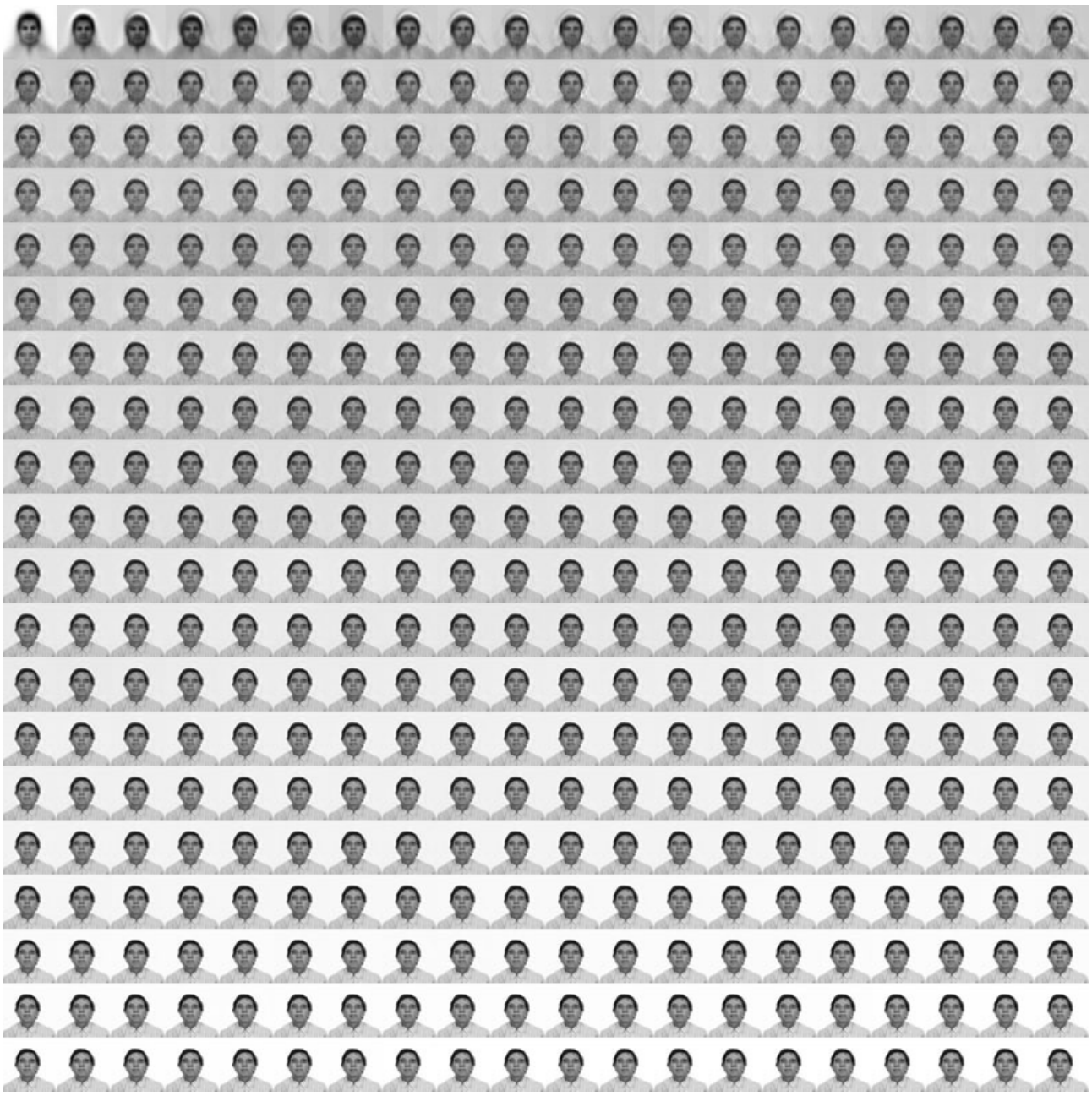


Figura 39 – Reconstrução de um retrato com K componentes (K=1:400)

Figura 40 – Reconstrução de um retrato 20, 40, 100, 200 e 400 componentes.

Fizemos também a experiência de reconstrução com uma imagem nova, que não pertencia à BD, e os resultados também são interessantes. Para este cálculo usámos um outro script Matlab que projecta a nova imagem no sistema de eixos correspondentes aos componentes PCA (vectores próprios), para calcular o vector de coeficientes PCA, o vector A, para fazer a reconstrução. A equação (50) calcula esses coeficientes; no fundo calcula o produto interno do vector imagem X

por cada um dos vectores do novo sistema de eixos ortogonal U. Cada produto interno por um eixo U_k corresponde à projecção de X nesse eixo.

$$A = X^T * U \qquad (50)$$

A reconstrução iterativa, para K componentes, é mais uma vez feita utilizando a equação (49), recorrendo ao vector A_K, construído da mesma forma a partir do vector A.

A Figura 41, a Figura 42 e a Figura 43 ilustram o erro na reconstrução e a reconstrução de um retrato, obtidos para K=1:400, para um retrato novo.

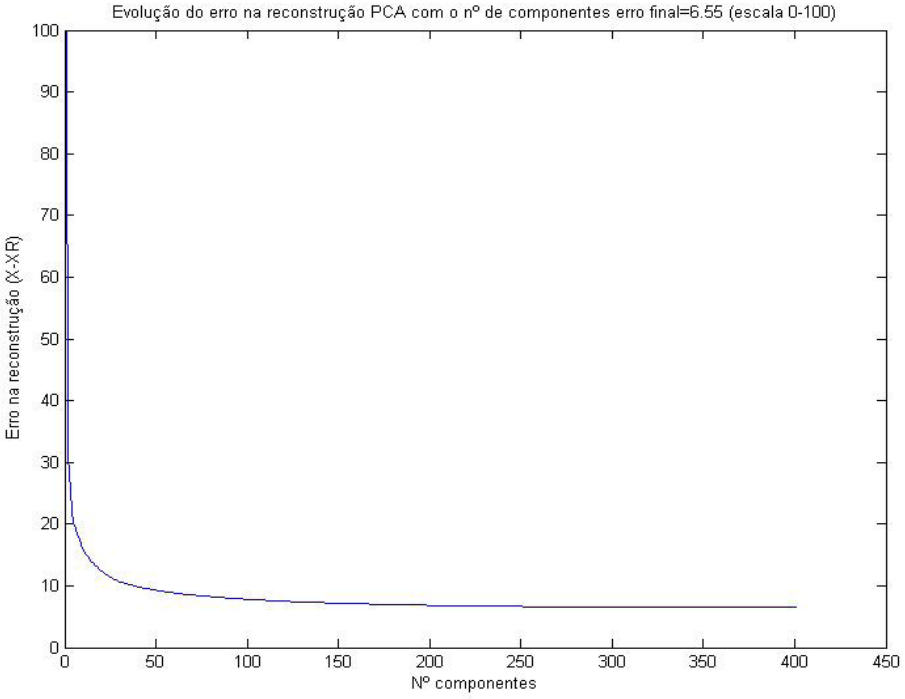


Figura 41 – Erro da reconstrução com K componentes (K=1:400)

De realçar o facto de se conseguir reconstruir um retrato novo, com um erro final a tender para cerca de 6% e mais ainda o facto do retrato reconstruído ser reconhecível com cerca de 60 componentes. Aliás a reconstrução pouco melhora com mais de 100 componentes, o que confirma a evolução do erro que se observa na Figura 41.

Ou seja:

1. Verificámos que apesar de a imagem não pertencer à BD, conseguimos reconstruir a imagem e conseguimos reconhecer a face, curiosamente, tal como no caso anterior, com pouco mais de 20 componentes.
2. Verificámos ainda que o erro diminui rapidamente mas neste caso não baixa de cerca de 6 numa escala de 0 a 100.

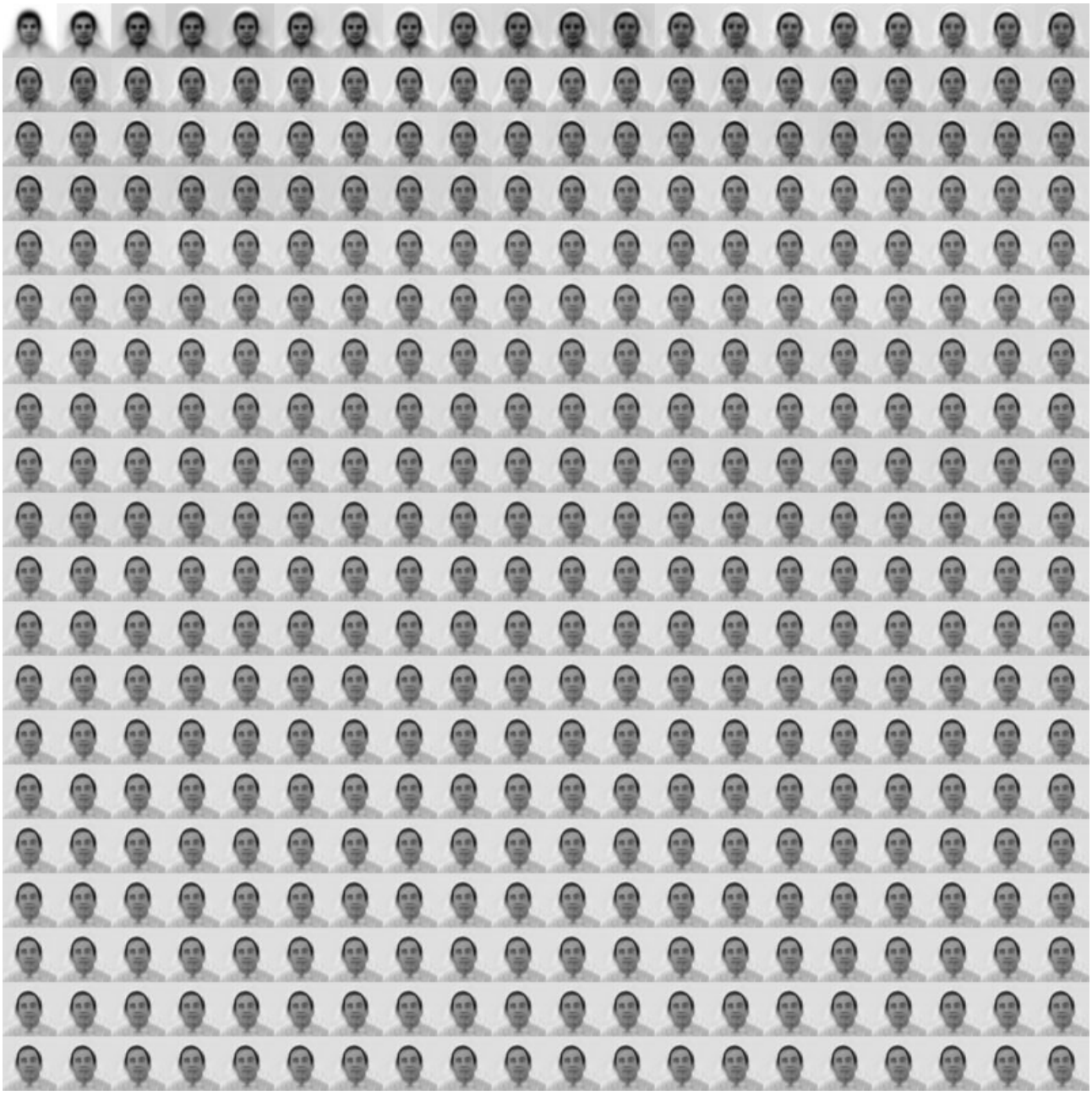


Figura 42 – Reconstrução de um retrato novo com K componentes (K=1:400) [6]

Figura 43 – Reconstrução de um retrato novo com 60, 100 e 400 componentes.

6 Ao longo deste trabalho utilizámos várias versões da BD, retirando ou introduzindo retratos, de acordo com os resultados que procurávamos. Para obter esta reconstrução com o retrato do André Puga, tivemos de substituí-lo na BD, colocá-lo no grupo de teste e recalcular o PCA. Ao fazer a edição final deste documento fomos confrontados com uma decisão: deveríamos manter aqui esta reconstrução ou substituí-la por outra, onde o peso da tristeza se não fizesse sentir? Decidimos manter a figura como metáfora de que acreditamos que é possível reconstruir e relembrar um rosto, mesmo não estando presente na BD de retratos quotidianos das nossas vidas. E poderíamos perguntar "Quantos componentes serão necessários para reconhecer o teu sorriso?"

4.2.1 Aprendizagem não-supervisionada

Depois de se construir uma BD de conhecimento de rostos, perante um novo rosto haverá duas opções: decompor e reconstruir esse rosto com base no conhecimento adquirido com a BD existente, como se fez na secção anterior – ver a Figura 42 e a
Figura 43 – ou porventura integrá-lo também na BD, o que obrigará a refazer o cálculo dos componentes, no nosso caso reaplicando o algoritmo PCA.

Seria de esperar que essa decisão pudesse ser tomada com base na quantidade de informação nova que o novo rosto contém – se fizermos a analogia com o que acontece no sistema neurológico estaremos em presença de um problema de aprendizagem não supervisionada de um novo rosto – podendo também ter em conta factores sociais, culturais e psicológicos.

Tentámos obter alguma informação sobre este processo com a nossa BD. Na BD que construímos para o nosso projecto existem muito poucos rostos de pessoas de cor ou de asiáticos[7], o que reflecte no fundo a pequena percentagem de pessoas com essas características na população da Universidade do Porto. Pensámos, por isso, que utilizando um retrato de uma pessoa com essas características nos testes, pudéssemos encontrar diferenças no processo de reconstrução, acreditando que esses retratos conteriam maior informação nova e por isso a reconstrução pudesse ser afectada. Para comparar resultados começámos por tentar reconstruir um rosto de um negro e de um asiático, que integrassem a BD de retratos, e depois, em alternativa, rostos com essas características que estivessem no grupo de testes, mas os resultados em nada se pareciam diferenciar dos resultados obtidos com outros elementos da BD. A Figura 44 mostra por exemplo a reconstrução de um rosto negro que fazia parte da BD, com 40, 100, 200 e 400 componentes e não há diferenças visíveis com o que se obtivera com outros retratos da BD. A curva que obtivemos para o erro na reconstrução foi também idêntica à da Figura 38.

Figura 44 – Reconstrução de um retrato 40, 100, 200 e 400 componentes

Perante este resultado decidimos levar o exercício ao limite. Experimentámos retirar da BD de retratos utilizada pelo PCA todos os rostos que de alguma forma se identificassem com essas fisionomias. Recalculámos os componentes PCA. Tentámos depois projectar e reconstruir vários rostos com essas características e surpreendentemente o resultado obtido em nada se diferenciou dos resultados obtidos com qualquer outro rosto que tentámos reconstruir com a mesma BD. No caso representado na Figura 45 o erro final é mesmo um pouco inferior a 6% e parece-nos também que se consegue reconhecer a identidade da pessoa do retrato. A Figura 45 e a Figura 46 apresentam a evolução do erro e a evolução no processo de reconstrução para o exemplo que apresentamos.

Este resultado poderia parecer surpreendente, pois um sistema de componentes, construído a partir de elementos de determinado tipo fisionómico, permite reconstruir com erro idêntico elementos do que consideraríamos ser outro tipo. Estávamos à espera de obter um resultado diferente para o erro, que pudesse servir de justificação para se decidir integrar na BD esse rosto

7 Temos consciência de que esta classificação é subjectiva e não científica e de que assenta muitas vezes em preconceitos sociais e culturais. Não pretendemos aprofundar este tema, pois o nosso objectivo é de apenas testar a reconstrução de rostos com características fisionómicas diferentes ou que nos parecem diferentes.

que, por parecer diferente, pudesse ter mais problemas na reconstrução e no reconhecimento. De facto, a nossa experiência diz-nos que temos maior dificuldade em memorizar e reconhecer rostos em algumas regiões do mundo que visitamos pela primeira vez e onde as características fisionómicas predominantes são diferentes.

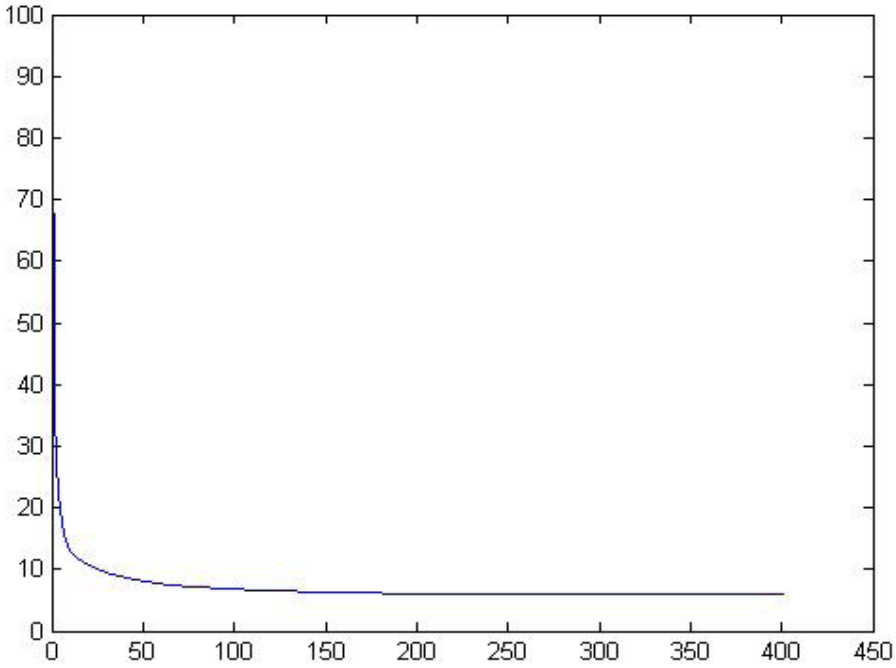


Figura 45 – Erro para retrato com fisionomia africana - à BD que não contém esse tipo de rostos

Este resultado, sendo interessante, pois indica que a diferença que esperávamos na reconstrução desses retratos não existe, pode não ser no entanto significativo do ponto de vista da tomada decisão, perante um novo rosto com essas características: decidir se se justifica ou não integrá-lo na BD, recalcular os componentes, ou seja aprender duradouramente com ele. Uma análise fundamentada obrigaria a um cálculo da entropia de comunicação do sistema existente, considerando a entropia da comunicação sem integração e a entropia resultante se integrássemos o novo retrato, perante um determinado universo de imagens que pretendemos reconhecer. Ou seja, sendo possível reconstruir o novo retrato e reconhecê-lo, como constatámos, isso pode implicar um aumento significativo da entropia nos canais de comunicação e portanto justificar a sua aprendizagem. Esta análise, fazendo parte do nosso horizonte, situa-se no entanto para já fora do horizonte temporal do mestrado, pois obrigaria a um cálculo alargado e pesado a partir dos retratos que integram a BD de conhecimento e de teste.

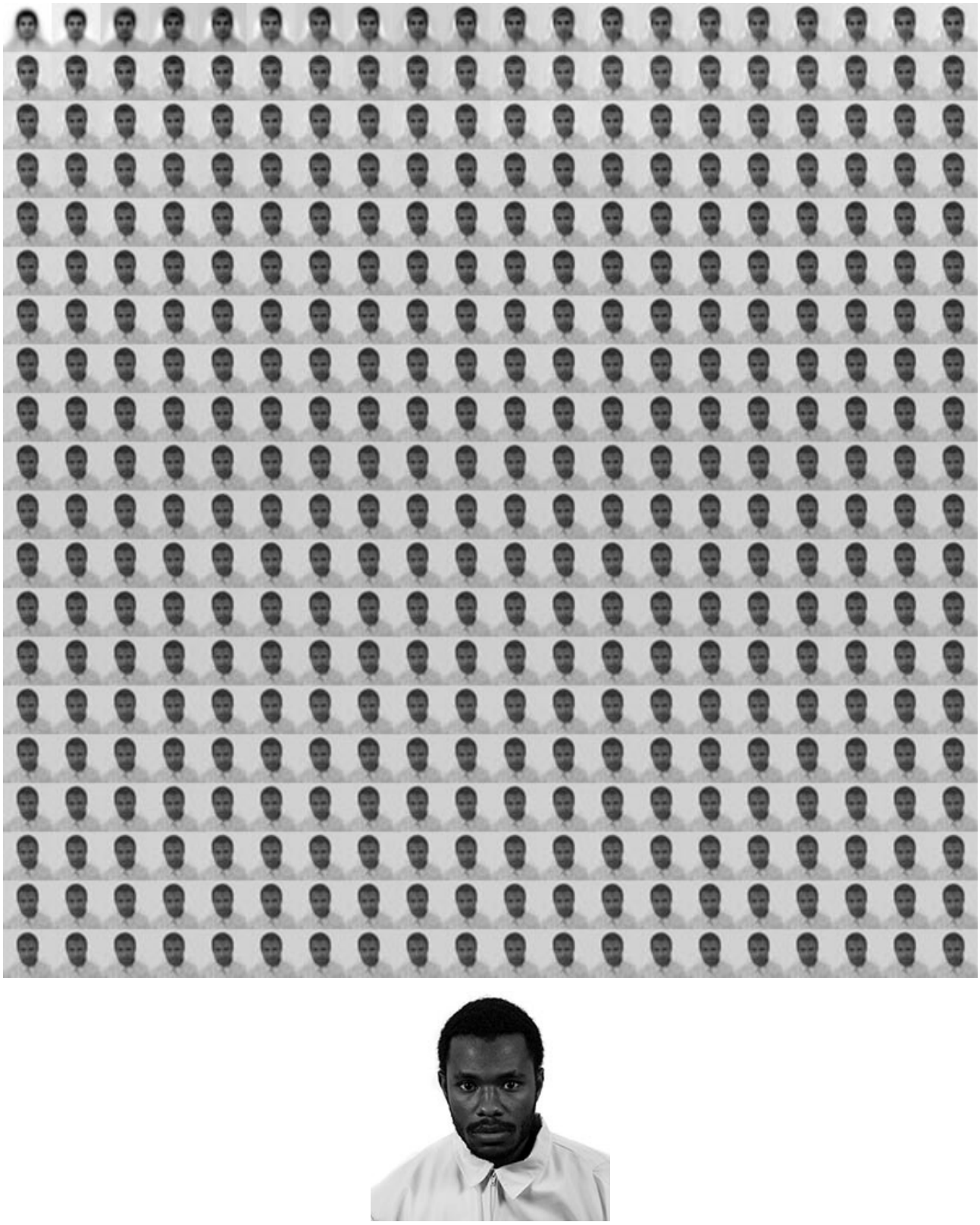
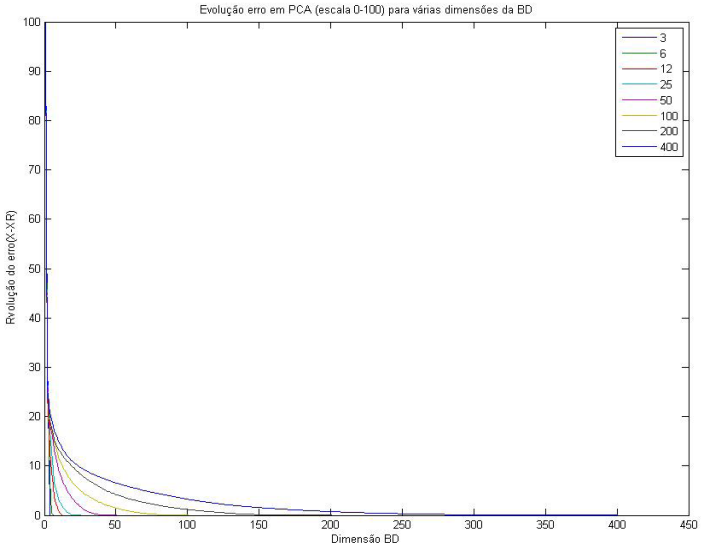


Figura 46 a e b – Reconstrução de um retrato de cor que não pertence à BD e o retrato original

4.2.2 Dimensão da BD para o PCA

Os resultados obtidos, em particular a evolução do erro muito idêntica na reconstrução de imagens, para duas BDs com dimensões diferentes[8], levou-nos a procurar analisar o comportamento do erro e da reconstrução de retratos, para BDs com dimensões diferentes. As Figura 47 a e b mostram:

- a evolução do erro na reconstrução do mesmo retrato, para BDs construídas com 400, 200, 100, 50, 25, 12, 6 e 3 retratos
- o nº de componentes necessários para reconstruir esse retrato com erro < 0.01 para cada uma dessas BDs



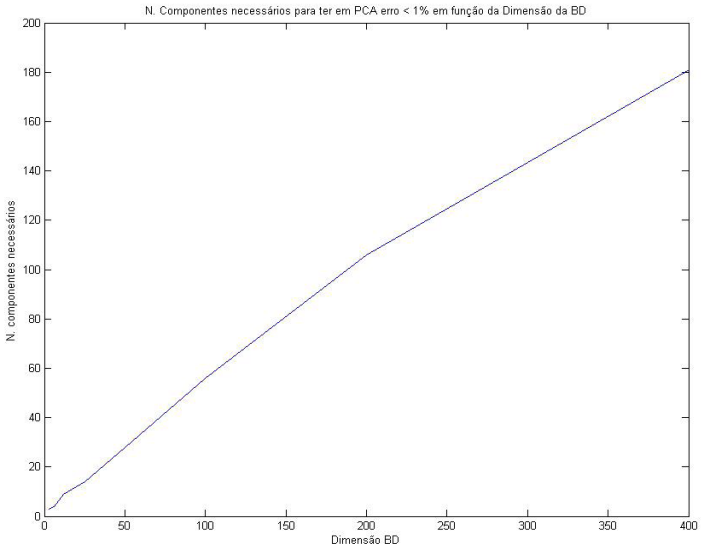


Figura 47 a e b – Resultados para BDs com dimensões diferentes

Observamos que:

- o erro final de reconstrução para um retrato da BD é sempre zero e diminui mais rapidamente quando a dimensão da BD baixa

8 Antes de termos construído a nossa BD trabalhámos com uma BD pública – a BD faces94 – com apenas 110 retratos

- as curvas de evolução do valor do erro quase coincidem para os primeiros componentes, até um erro de cerca de 20%, que se atinge com muito poucos componentes
- o nº de componentes necessários para garantir um determinado erro varia quase linearmente com a dimensão

Estes resutados podem surpreender mas são de alguma forma triviais e apenas confirmam a capacidade do PCA, como método, para reconstruir as imagens sem erro, a partir dos componentes ortogonais que calcula, qualquer que seja a dimensão dos dados[9].

Mas será interessante testar a forma como a dimensão da BD afecta a reconstrução de retratos novos e comparar o valor final para o erro de reconstrução de um retrato novo. A Figura 48 mostra a evolução do erro final de reconstrução de um retrato novo com a dimensão da BD. O erro final de cerca de 6, que obtivemos com a nossa BD de 400 retratos, aumenta claramente, se o número de retratos da BD diminuir. Se tivermos em conta os resultados da secção 6, onde demonstramos que a zona crítica para o reconhecimento parece situar-se num erro > 9, poderemos concluir que neste caso - retratos com 200x200 pixéis - necessitamos de um número de imagens da BD de cerca de 170 retratos (valor estimado) para assegurar um erro de reconstrução inferior a 9.

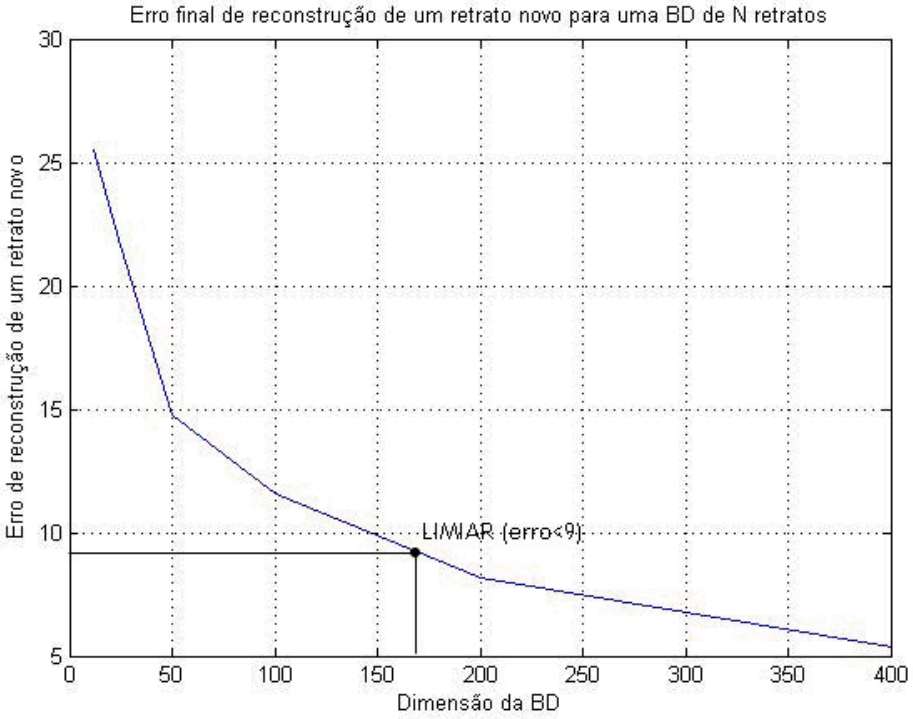


Figura 48 – Erro final de reconstrução de um retrato novo para uma BD de N retratos

170 vai ser portanto o limiar da dimensão da BD para conseguir reconstruir retratos novos reconhecíveis com o PCA, para retratos com uma resolução de 200x200 pixéis. Notar que apesar de necessitarmos de 170 retratos na BD, em rigor a reconstrução do retrato novo com um erro < 9 necessita de muito menos componentes, devido ao andamento das curvas de erro de reconstrução. Observa-se na Figura 49 que por exemplo neste cenário, com uma BD de 170 retratos, bastariam menos de 70 componentes para garantir um erro < 9.

9 Este resultado, podendo ser surpreendente no domínio das imagens, é fácil de demonstrar no domínio da matemática, pois o que PCA faz basicamente é uma mudança de coordenadas para os vectores imagem, permitindo por isso a sua reconstrução no novo sistema de eixos.

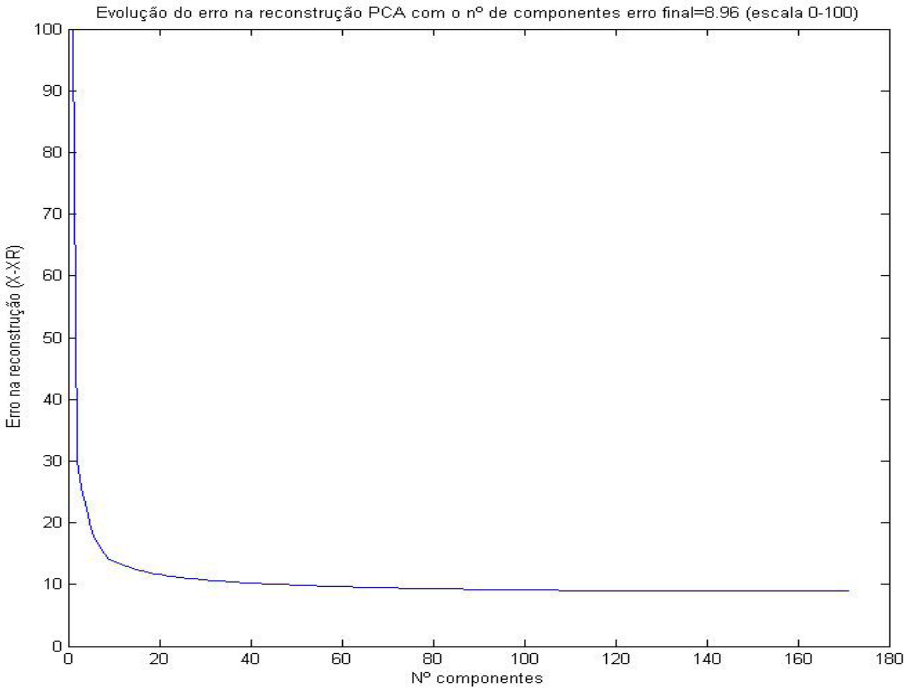


Figura 49 – Erro final de reconstrução de um retrato novo para uma BD de 170 retratos

Poderemos também analisar a evolução deste limiar com a resolução dos retratos, comparando o valor do limiar para BDs com 100x100, 50x50 ou 20x20 pixéis. A Figura 50 mostra os resultados obtidos.

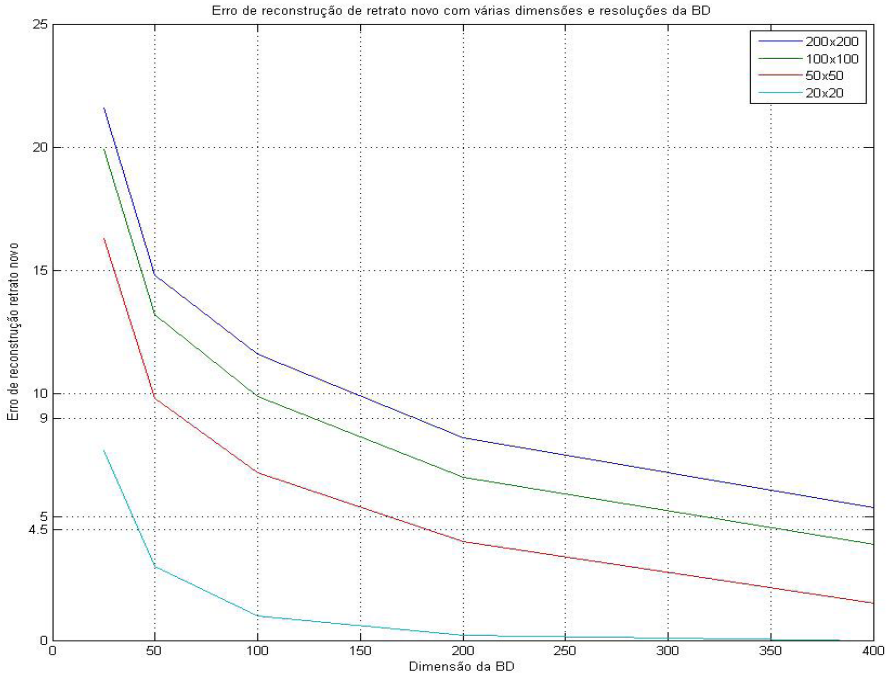


Figura 50 – Erro final de reconstrução de um retrato novo
BD de N retratos com diferentes resoluções – 20x20, 50x50, 100x100 e 200x200

Obtemos curvas com andamentos idênticos para o erro de reconstrução, para as diferentes resoluções. De notar que com 20x20 pixéis e com 400 retratos qualquer novo retrato é reconstruído sem erro, pois o número de retratos na BD é igual à resolução.

Este resultado é importante pois permite estimar as dimensões necessárias para as BDs, em função da resolução dos retratos utilizados.

4.3 Resultados com o ICA

Desenvolveram-se um conjunto de scripts em MATLAB, equivalentes aos scipts que descrevemos para o PCA, utilizando o algoritmo *fastICA*, para fazer a decomposição da matriz de vectores de imagens em componentes ICA e para avaliar a reconstrução dos retratos a partir desses componentes.

Mas o algoritmo *fastICA* oferece um conjunto de quatro funções distintas, para procurar esses componentes. Essas funções são:

pow3	g(u)=u^3
gauss	g(u)=u*exp(-a2*u^2/2)
skew	g(u)=u^2
tanh	g(u)=tanh(a1*u)

O algoritmo utiliza por defeito, para induzir a convergência para soluções não gausseanas, a função *pow3*.

4.3.1 Comparação entre as várias funções utilizadas no fastICA

As secções seguintes comparam os primeiros resultados obtidos com o *fastICA* para cada uma das quatro funções, com base em diferentes itens. Esta comparação foi realizada para decidir qual a função em que o nosso trabalho se focaria.

Os componentes ICA, para cada uma das quatro funções que o algoritmo *fastICA* disponibiliza, foram calculados com um pequeno script de Matlab, bastante parecido com o script que foi utilizado no PCA para calcular os componentes PCA.

Esse script lê todos os retratos da BD a partir de uma directoria de imagens, transforma-os em vectores, retira-lhes o valor médio e normaliza-os com comprimento unitário, para construir a matriz X de retratos. Depois aplica o algoritmo *fastICA* e calcula as matrizes A, e U, onde cada vector U_i neste caso representa uma componente ICA, e onde a matriz A é a matriz de mistura com os coeficientes ICA. Depois, para cada componente ICA, calcula a matriz imagem associada e guarda-as numa directoria de componentes em formato tiff ou jpg. As matrizes A, e U são guardadas num ficheiro de dados Matlab.

4.3.1.1 Eficiência do algoritmo

Para todas as funções disponibilizadas, o algoritmo *fastICA*[10] convergiu a partir da nossa BD de retratos e calculou um a um todos os componentes ICA, calculados por um processo iterativo, em

10 Para cada uma das 4 funções se correu o algoritmo pedindo o cálculo de 400 componentes.

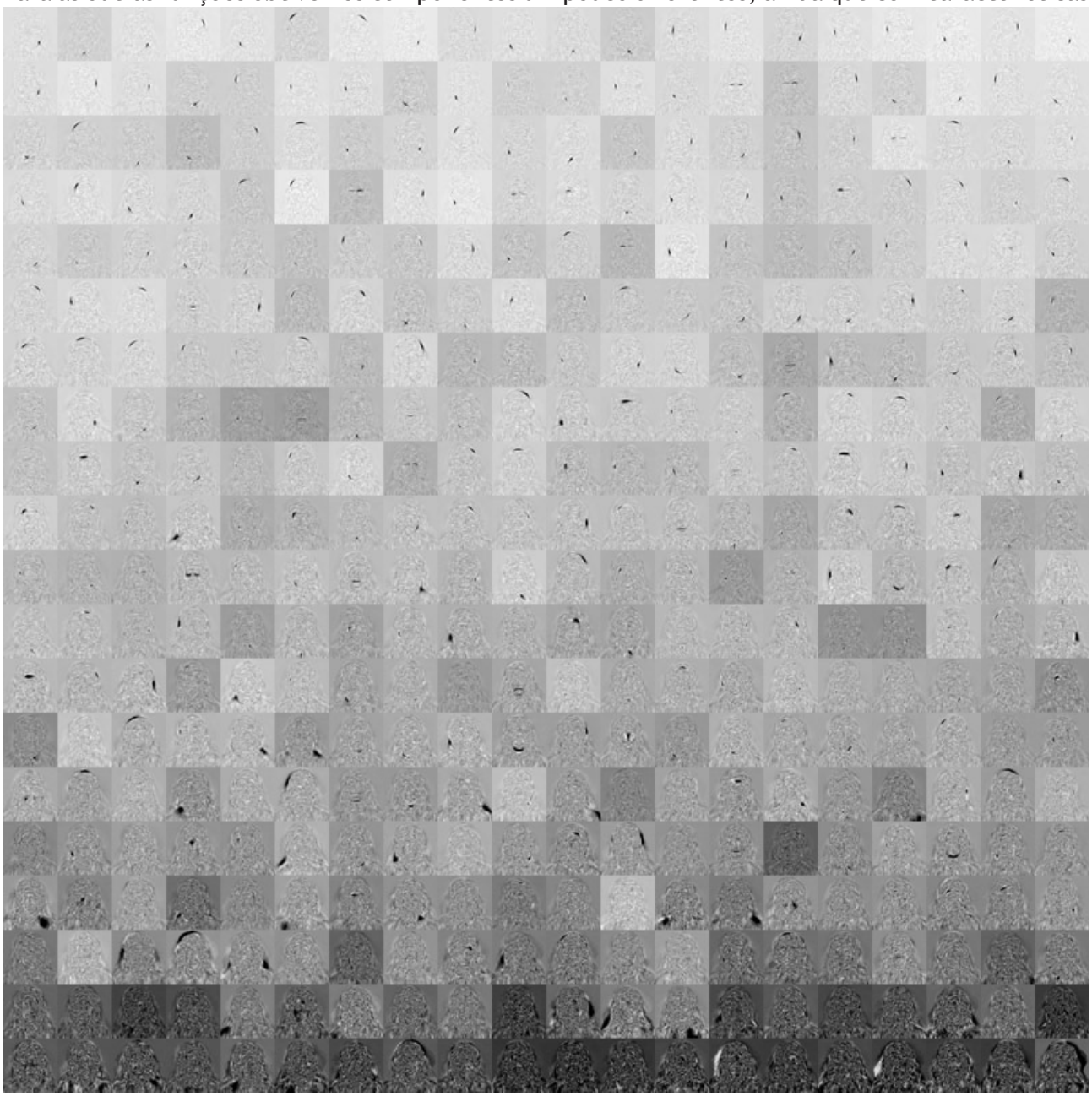
que o número de iterações por componente depende grandemente da função escolhida. Os 400 componentes foram calculados pelo algoritmo para cada uma das quatro funções, excepto para a função *skew,* onde alguns componentes não convergiram, e por isso o algoritmo parou depois de calcular apenas 382 componentes.

Apesar de todas as funções permitirem a convergência do algoritmo e terem como resultado o cálculo dos componentes ICA, observou-se que as funções *pow3* e *skew* são bastante mais rápidas, permitindo o cálculo de todos os componentes em muito menos tempo.

4.3.1.2 Componentes ICA

A Figura 51 apresenta os componentes ICA obtidos para a nossa BD com a função utilizada por defeito pelo algoritmo, a função *pow3*.

Para as outras funções obtivemos componentes um pouco diferentes, ainda que com características



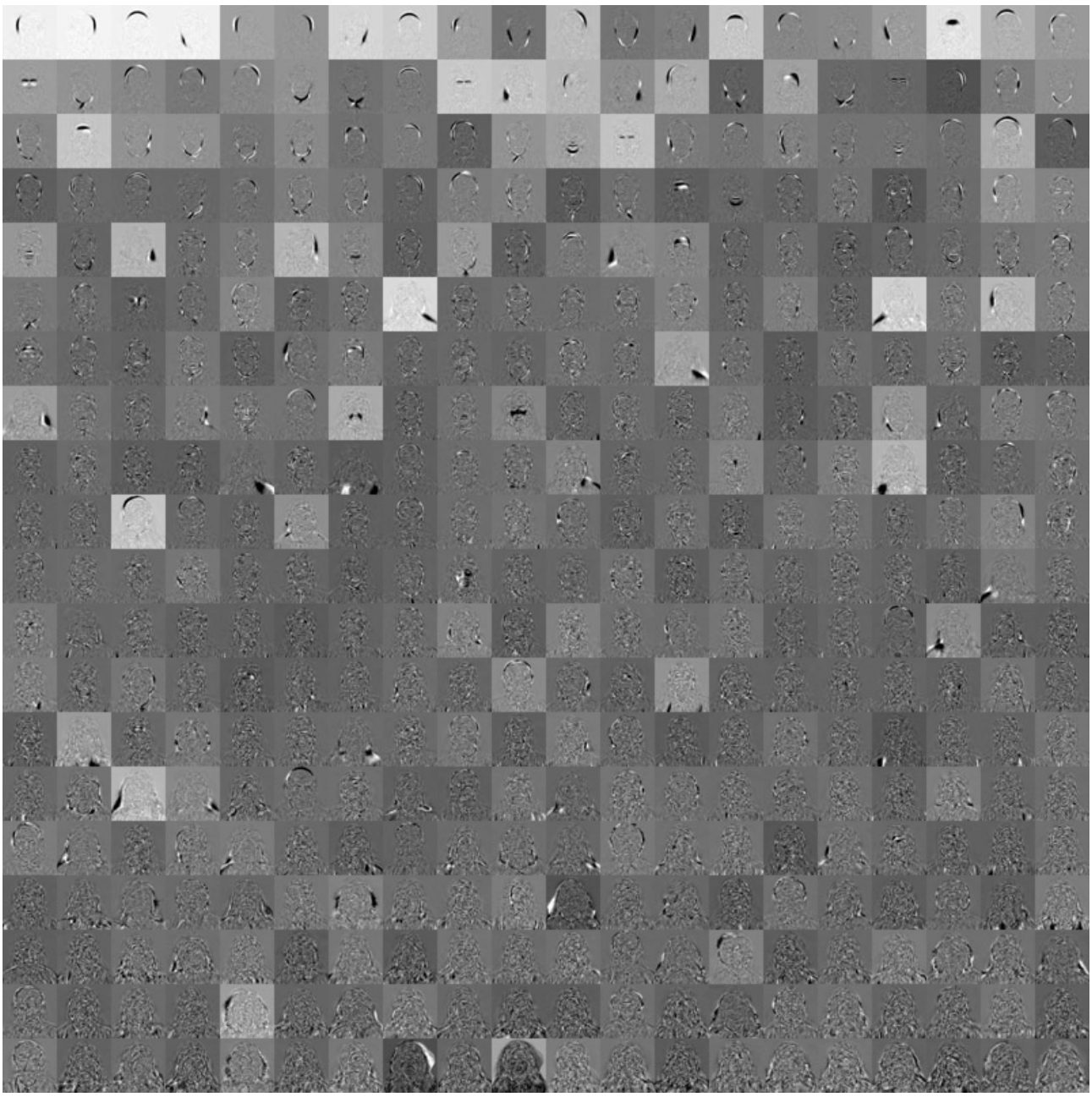


Figura 52 – Componentes ICA com a função de *gauss*

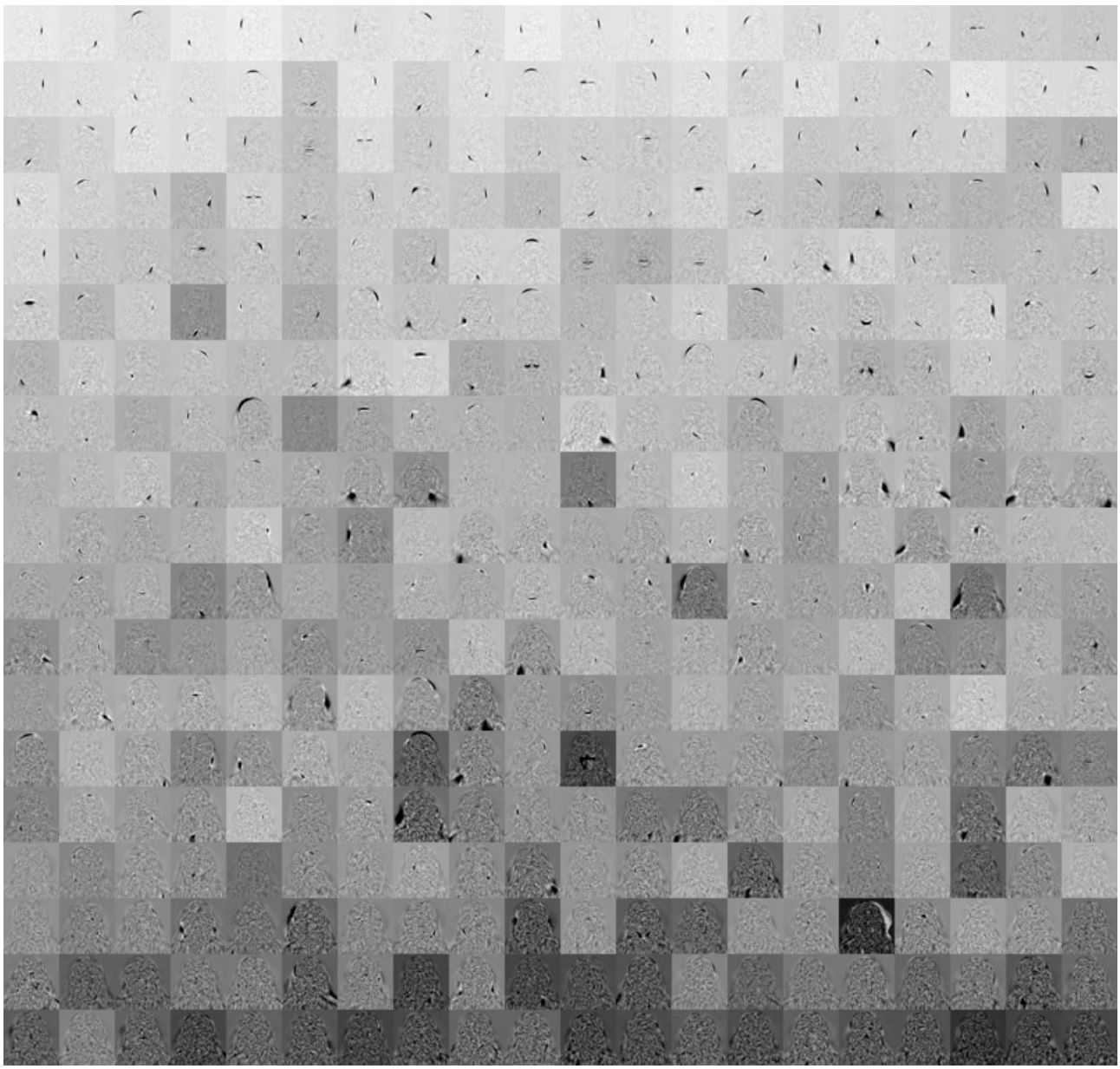


Figura 53 – Componentes ICA com a função de *skew*

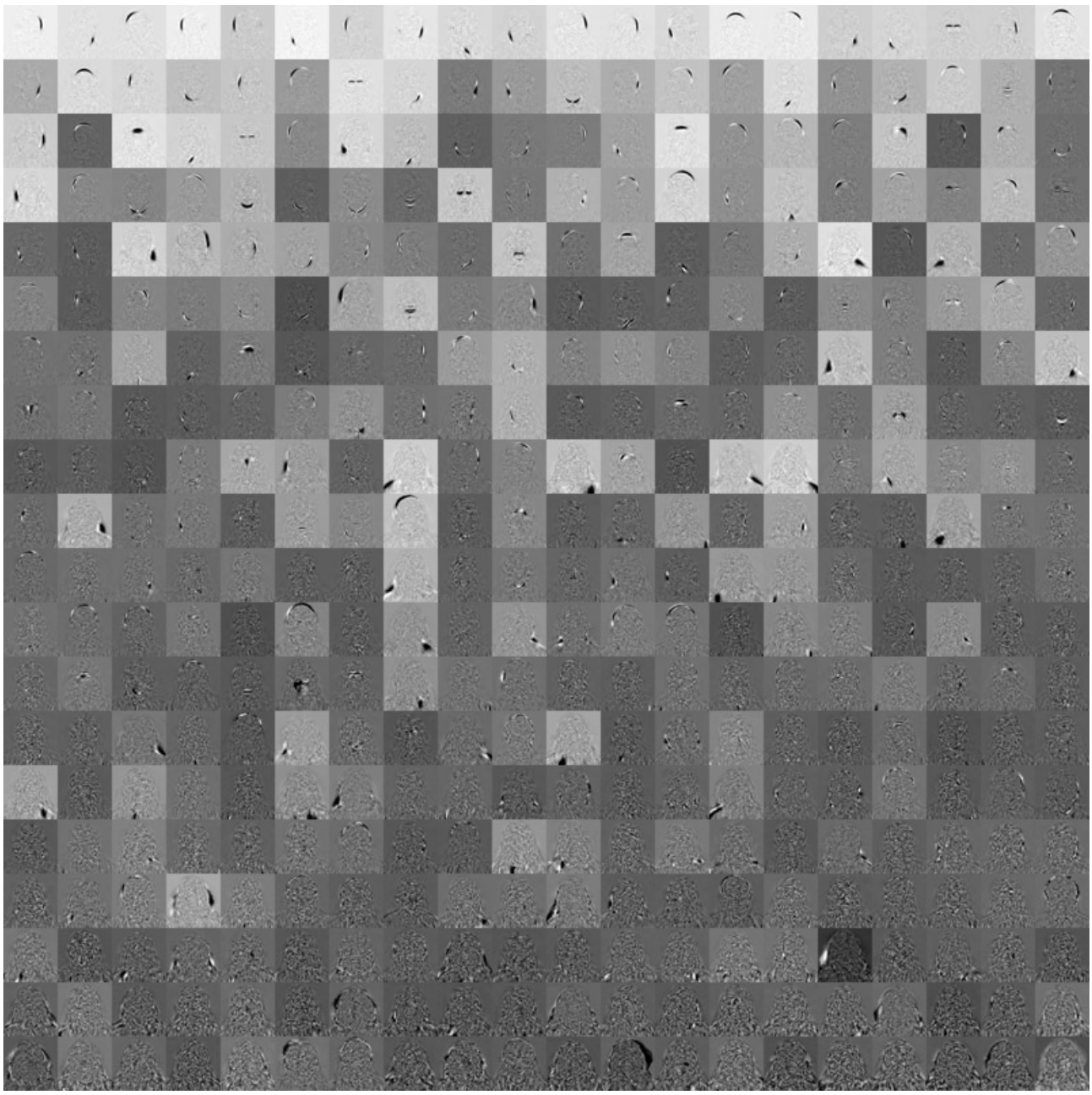


Figura 54 – Componentes ICA com a função de *tanh*

À primeira vista todos os componentes ICA que obtivemos parecem também corresponder a imagens fantasmagóricas de toda a face, no entanto observam-se pequenas manchas numa região específica do rosto para cada componente: na boca, nos olhos, no nariz, nas sobrancelhas, etc. O algoritmo *pow3* parece definir, a olho nu, zonas mais pequenas. Pretendíamos confirmar se esta observação poderia corresponder à ideia de que o ICA, ao definir critérios específicos, associados ao carácter não-gaussiano dos dados, identifica atributos físicos e visíveis das imagens e poderia portanto ser em certa medida complementar do PCA, método que corresponderia mais a um conhecimento genérico e global dos rostos.

Temos de reconhecer que os resultados que obtivemos não foram conclusivos sobre se os componentes ICA permitem identificar zonas e atributos específicos dos rostos, mas obtivemos um conjunto significativo de resultados, que apresentamos ao longo deste trabalho, que parecem reflectir a complementaridade entre o PCA e o ICA, e que poderão estar associados a essa característica do ICA.

4.3.1.3 O erro na reconstrução no ICA

Apresentamos a seguir a evolução do erro na reconstrução de uma mesma imagem, para cada uma das quatro funções.

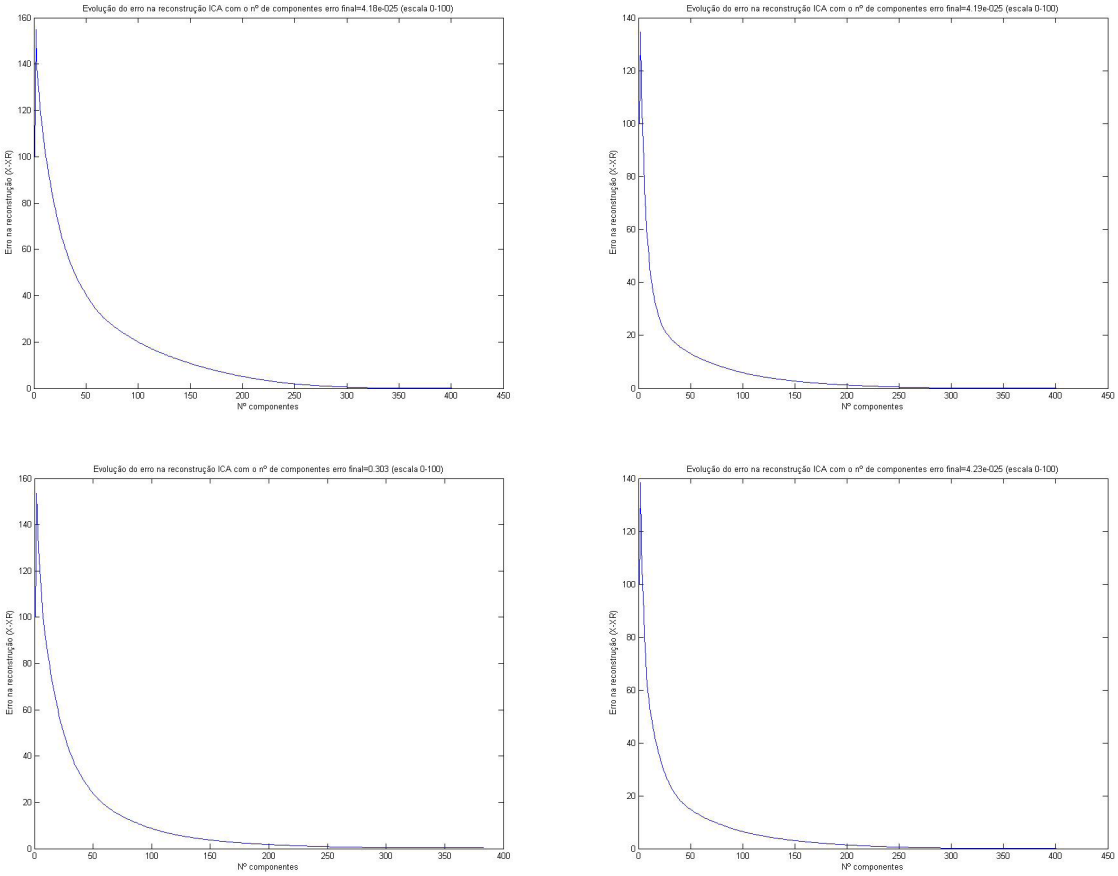


Figura 55 – Erro na reconstrução da mesma imagem com ICA pow3, gauss, skew, tanh

O andamento do erro para a reconstrução da mesma imagem é semelhante para as 4 funções, mas a convergência para zero é mais ou menos rápida. Observa-se que o erro ultrapassa ligeiramente 1 (100%) na primeira iteração, ao contrário do PCA, o que não surpreende pois, ao contrário do PCA, o ICA não está preso às direcções de variação máxima, mas depois converge de uma forma rápida e monótona para zero. O valor do erro final (usando todos os componentes) é zero, obtêm-se valores da ordem de 1 e-025 em 100, excepto para a função *skew*, onde o valor final é de 0.0003 em 100, mas onde também o algoritmo só conseguiu encontrar 382 componentes em vez dos 400 encontrados com as outras 3 funções.

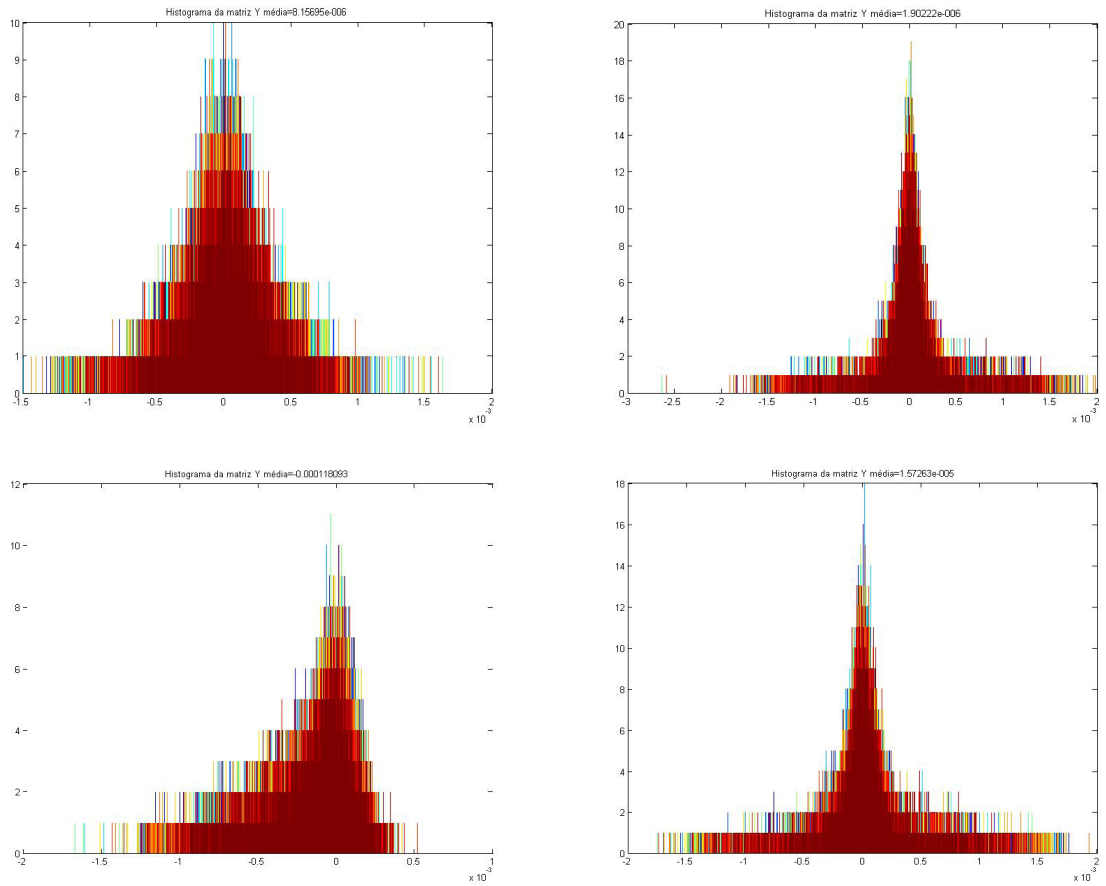
4.3.1.4 Avaliação das estatísticas dos coeficientes no ICA

Pretendíamos avaliar qual a função que nos interessava mais utilizar no nosso trabalho, das quatro que o *fastICA* oferece. As funções utilizadas pelo ICA permitem procurar o carácter não-gaussiano das variáveis, com base em critérios diferentes, podendo nomeadamente ter como resultado estreitar ou alargar a função densidade de probabilidade, relativamente à distribuição normal. Em princípio no nosso caso interessam-nos as soluções que estreitam mais essa função. Interessar-nos-á uma função que conduza a componentes ICA tão esparsas quanto possível, ou seja que permitam reduzir o número de componentes, para garantir um determinado erro. Existem duas estatísticas que permitem traduzir a não gausseanidade de uma variável aleatória: a estatística de assimetria (*skewness)* que mede a simetria dos dados, ou melhor a não-simetria dos dados e a estatística de

‘achatamento’ (*kurtosis)* que mede a maior ou menor centralidade dos dados, ou seja se a fdp é mais ou menos “achatada” no valor médio. A estatística de assimetria é simplesmente o momento normalizado de 3ª ordem das variáveis aleatórias e a estatística de ‘achatamento’ o momento normalizado de 4ª ordem das variáveis aleatórias. A estatística de ‘achatamento’ permite também, de alguma forma, medir o carácter não gaussiano das variáveis: os valores elevados positivos correspondem funções de distribuição com um pico muito acentuado no valor médio. Portanto a estatística de ‘achatamento’ permite de alguma forma também medir a ‘esparsidade’ dos dados. Uma forma de analisar estas estatísticas é através de histogramas. A Figura 56 apresenta os histogramas obtidos por um pequeno script em Matlab, utilizando os coeficientes calculados para cada uma das quatro funções que o *fastICA* oferece: *pow3*, *gauss*, *skew* e *tanh*, por esta ordem. Para cada função teremos 3 histogramas:

- o histograma para os coeficientes A_i, resultantes da decomposição ICA realizada
- o histograma dos momentos normalizados de 3ª ordem (*skewness*)
- o histograma dos momentos normalizados de 4ª ordem (*kurtosis*)

O script calcula também os valores médios das estatísticas, obtidos para o conjunto dos 400 retratos da BD.



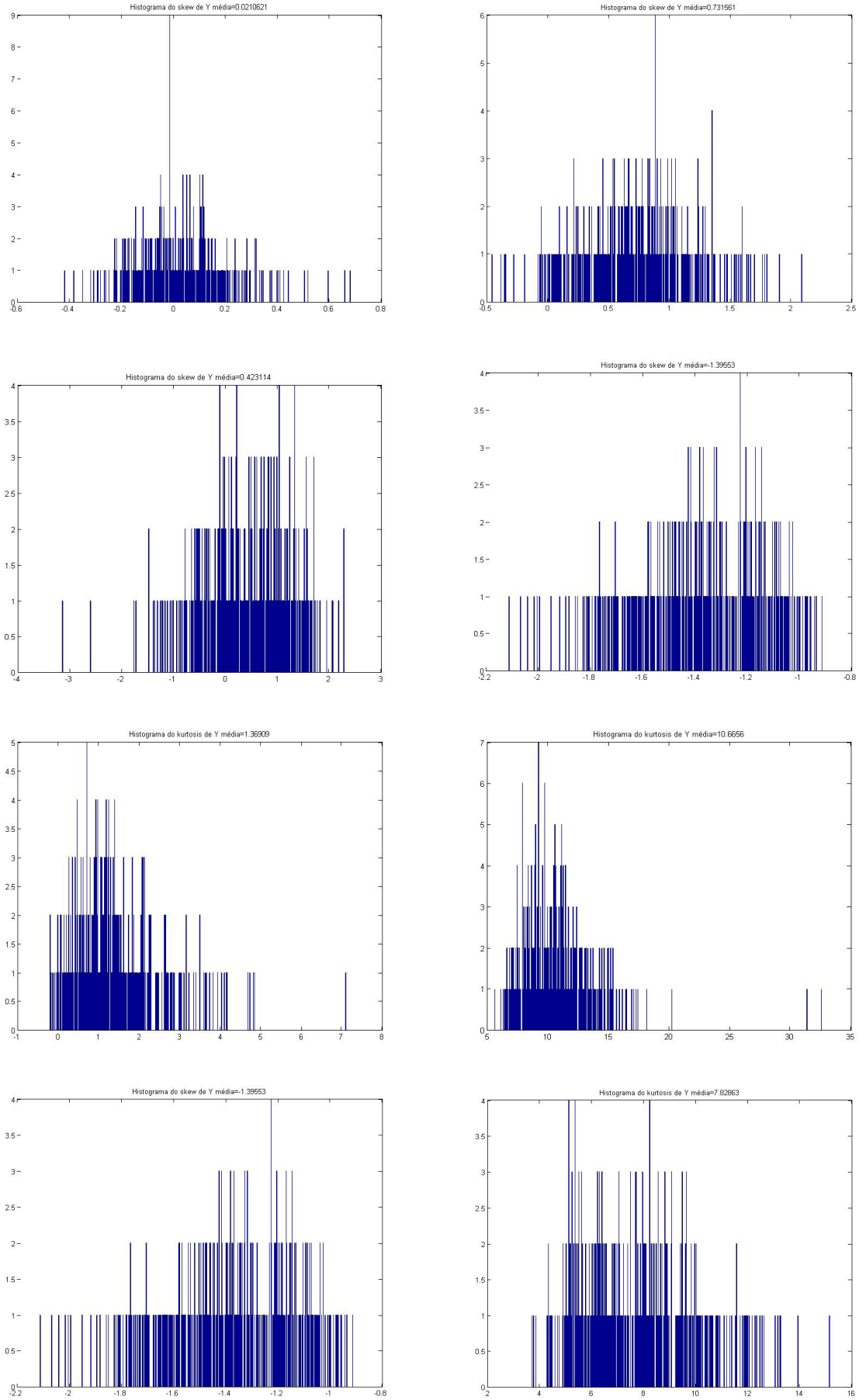


Figura 56 – Conjunto de histogramas para cada uma das funções: pow3, gauss, skew e tanh

Observa-se que a função não linear que induz maior valor médio de *kurtosis* (K=10.7) para os vários Ai é a função de *gauss*. Este resultado coincide com o da Figura 55, onde se observa a evolução do erro. Observa-se que no caso de *gauss* se consegue um erro < E com menos componentes, devido precisamente ao carácter mais esparso dos dados. Observa-se ainda que o melhor resultado, a seguir, se obtém com a função *tanh* e que *pow3* conduz ao pior resultado de todos.

Com base neste resultado decidimos utilizar no nosso trabalho para o método ICA o algoritmo *fastICA* com a função de gauss.

4.3.2 A reconstrução de uma imagem com ICA

A reconstrução de uma imagem utilizando um sistema ICA oferece algumas dificuldades em virtude de se tratar de um sistema não ortogonal. Não é por isso possível projectar um vector nos eixos ICA e tentar reconstruí-lo com esses coeficientes, tal como se fez no PCA, para todas as imagens, pertencessem ou não à BD de conhecimento.

No caso de uma imagem que pertence à BD, pode-se reconstruí-la utilizando os coeficientes que o próprio algoritmo *fastICA* calcula (matriz A). Utilizámos um script Matlab, parecido com o que se utilizou no PCA. O script lê o ficheiro com o retrato a testar, vectoriza-o e normaliza-o, retirando a média e dando-lhe comprimento unitário. Utilizando os coeficientes ICA da matriz A faz uma reconstrução do retrato por um processo iterativo, utilizando de cada vez mais um componente de 1:400 componentes. Para isso, tal como se fez no PCA, é utilizado um vector de coeficientes A_K, obtido a partir do vector dos coeficientes ICA para esse retrato A_i – A_i corresponde à linha de ordem i da matriz A, a matriz dos coeficientes calculada pelo ICA, se o retrato for o retrato ordem i da BD. A_K é facilmente obtido a partir de A_i, em cada iteração K, fazendo zero todos os coeficientes de ordem> K, garantindo assim mais uma vez que na equação (49) que os componentes de ordem> K não são utilizados na reconstrução. É calculado mais uma vez o erro da reconstrução comparando X_R com X em cada iteração, utilizando a equação (48); cada vector retrato X_R assim reconstruído é recalculado como uma matriz imagem de 200x200 pixéis e guardado numa directoria de reconstruções com o formato tiff ou jpg.

A Figura 57 mostra o resultado obtido para o erro com o *fastICA* (método gauss), fazendo a reconstrução com os coeficientes A_i calculados pelo *fastICA*. A figura mostra a evolução do erro da reconstrução, comparando-o com o erro no PCA. A evolução do erro não é muito diferente do PCA. Observamos um erro que, tal como no PCA, tende claramente para zero, ainda que no PCA a curva convirja claramente mais rapidamente. O valor final do erro para a reconstrução, calculado para o ICA, foi cerca de 3.12 e-025 (idêntico ao PCA), ou seja, tal como no PCA, zero.

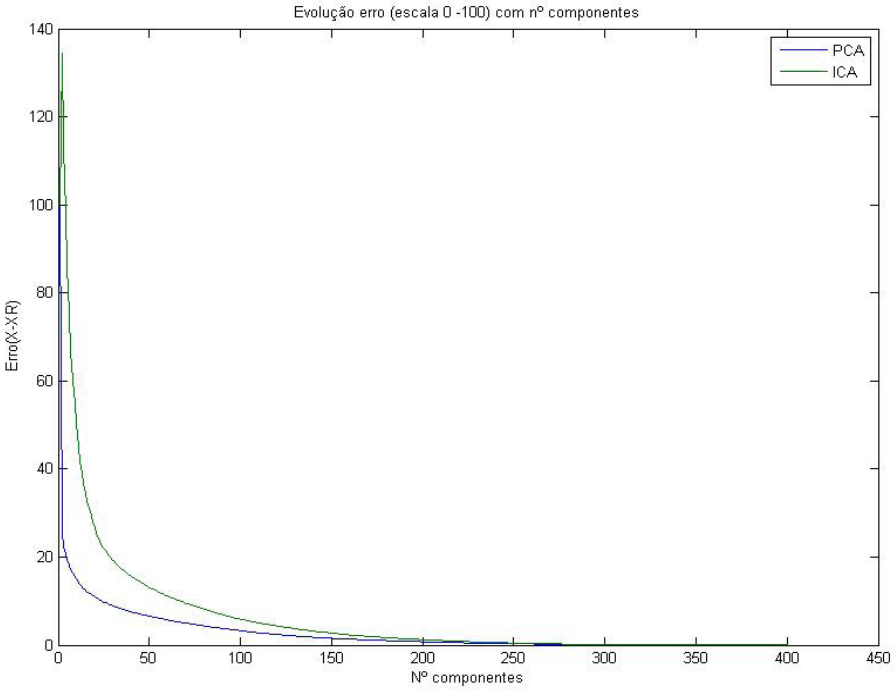


Figura 57 – Evolução do erro de reconstrução ICA e PCA com K=1 a 400

A Figura 58 mostra a evolução da imagem reconstruída com o nº de componentes ICA. A imagem reconstruída com o sistema ICA, tal como acontecia com o PCA, evolui também para a imagem original, ainda que tal como no erro a convergência seja mais lenta para os primeiros componentes. Se fizermos uma reconstrução de um retrato com apenas 50 componentes PCA, verificamos que precisamos de quase 100 componentes ICA para ter uma qualidade equivalente.

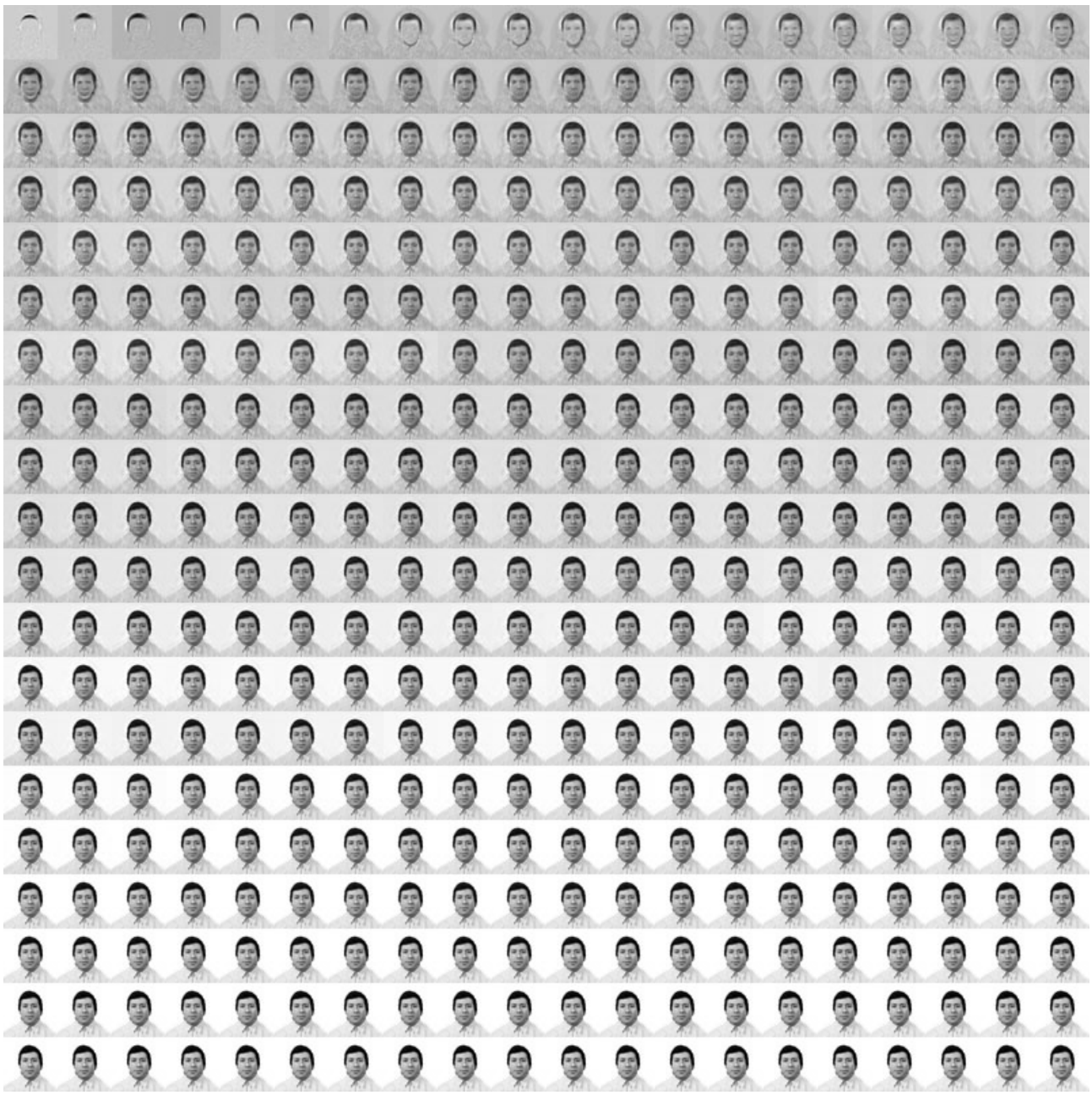


Figura 58 – Reconstrução de um retrato com K=1 a 400

Como utilizar o ICA para retratos que não pertencem à BD? Em virtude de os eixos ICA não serem ortogonais, uma imagem não pode ser decomposta nesse sistema por projecção, como fizemos no PCA, o que dificulta a reconstrução de retratos que não pertençam à BD com ICA. Vejamos porquê, utilizando como exemplo um sistema simples 3D, representado na Figura 59.

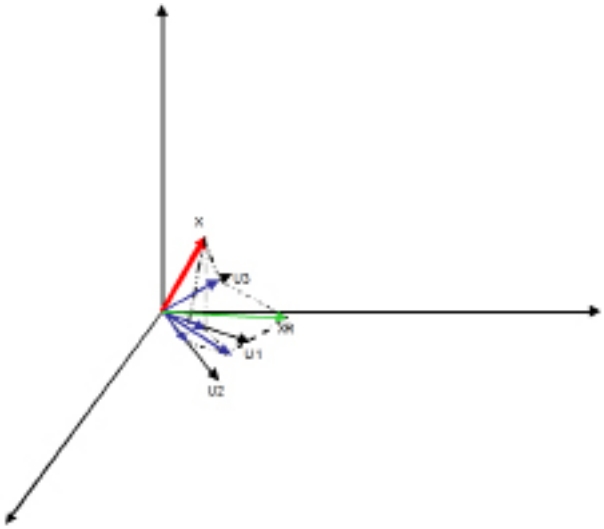


Figura 59 – Reconstrução de X num espaço 3D não ortogonal

Neste exemplo temos um sistema tridimensional e pretende-se decompor o vector X no espaço 3D em 3 componentes, U1, U2 e U3, que não são ortogonais. Se utilizarmos projecções obtemos X_R, muito longe de X. É fácil observar que teríamos uma melhor aproximação se apenas se utilizasse a componente U3 – *sparse coding*. É fácil no entanto observar que, se utilizarmos um novo sistema ortogonal, tal como no PCA, a projecção conduz sempre à recomposição do vector, que é única, havendo apenas uma mudança de coordenadas de X,Y,X para X',Y',Z'.

Para conseguir reconstruir qualquer imagem em qualquer sistema de eixos, ortogonal ou não, tivemos portanto de construir um algoritmo genérico que, por um processo iterativo, "projecta" qualquer vector, ou seja neste caso um vector imagem, no sistema de eixos utilizados, permitindo assim decompô-lo e reconstrui-lo de uma forma correcta.

O processo em que o algoritmo se baseia é iterativo e converge claramente, como se observa na Figura 60, onde se representa o erro na reconstrução, calculado para um retrato da BD, com o sistema de eixos ICA (*gauss*), utilizando esse algoritmo, a que chamamos algoritmo *pesqproj*, e onde se compara esse erro com o que foi calculado fazendo a reconstrução do mesmo retrato com os coeficientes do *fastICA*

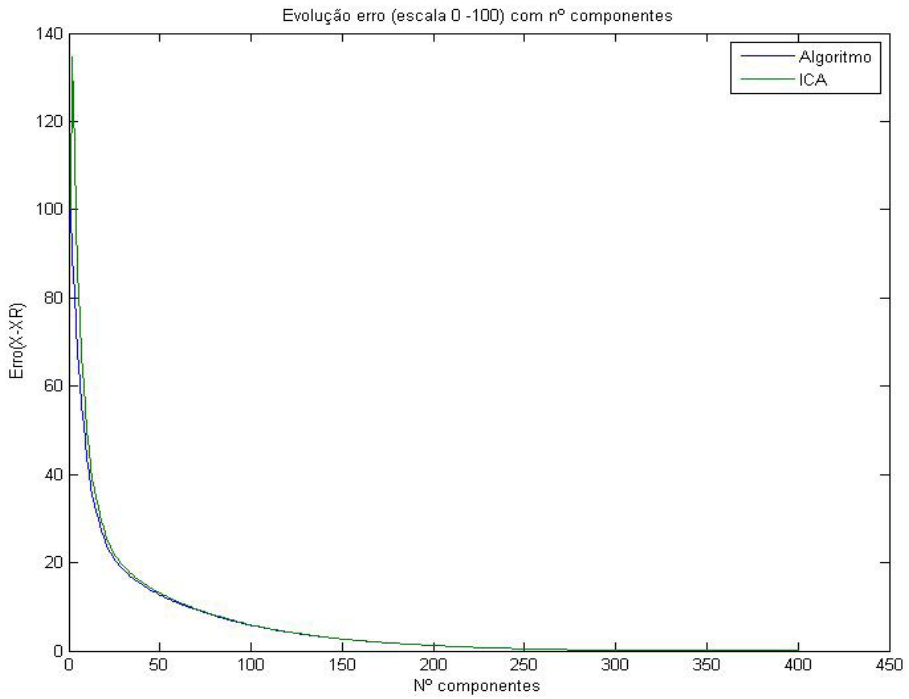


Figura 60 – Erro de reconstrução com coeficientes *fastICA* e com o algoritmo *pesqproj*

Verifica-se na Figura 60 que se obteve uma evolução para o erro idêntica, quase coincidente à que já tínhamos calculado com os coeficientes *fastICA*, o que permite validar o algoritmo.

4.3.3 O nosso algoritmo – o algoritmo *pesqproj*

O objectivo é decompor um qualquer vector X num qualquer sistema de eixos, ortogonal ou não, de forma a conseguir reconstruir esse vector, com os coeficientes obtidos. Este algoritmo é um algoritmo genérico, que iremos utilizar e adaptar em diversas situações e não só para o ICA. O ICA foi apenas a primeira situação com que fomos confrontados com a necessidade de um algoritmo com estas características.

X_R, o vector reconstruído, é inicializado a 0 e o vector erro inicial a X. O algoritmo, a que chamámos *pesqproj*, projecta o vector erro (o valor inicial é X como dissemos) no sistema de eixos AXIS a utilizar como referencial

$$proj = erro^T * AXIS \qquad (51)$$

A projecção obtém-se pelo produto interno do vector erro por cada um dos eixos AXIS – equação (51).

O resultado é um vector proj que contém os coeficientes desta projecção; procura-se o coeficiente de projecção $proj_i$ com valor máximo e escolhe-se essa projecção. O vector X reconstruído (X_R) e o vector erro que resultam de cada iteração vão ser

$$X_R = X_R + proj_i * AXIS_i \qquad (52)$$
$$erro = X - X_R \qquad (53)$$

O algoritmo é iterativo: projecta em cada iteração, o erro resultante da iteração anterior, nos

eixos do sistema, e escolhe sempre a projecção máxima; calcula novamente o erro e prossegue, até o erro final ser menor do que um limiar imposto, ou o nº de iterações ultrapassar um valor máximo. No final obtém-se uma matriz A de coeficientes de decomposição do retrato X no sistema de eixos utilizado.

A grande vantagem do algoritmo desenvolvido é de por um lado poder usar sistemas não ortogonais, como o *fastICA*, para imagens que não estão na BD, e é por outro lado permitir criar modelos híbridos que associam eixos PCA e eixos ICA ou NMF e poder decompor imagens nestes sistemas híbridos, como pretendemos fazer.

Utilizando o algoritmo *pesproj*() e os componentes ICA obtidos com a função de *gauss*, que identificámos como a mais favorável, reconstruímos então uma imagem que não pertence à BD de conhecimento.

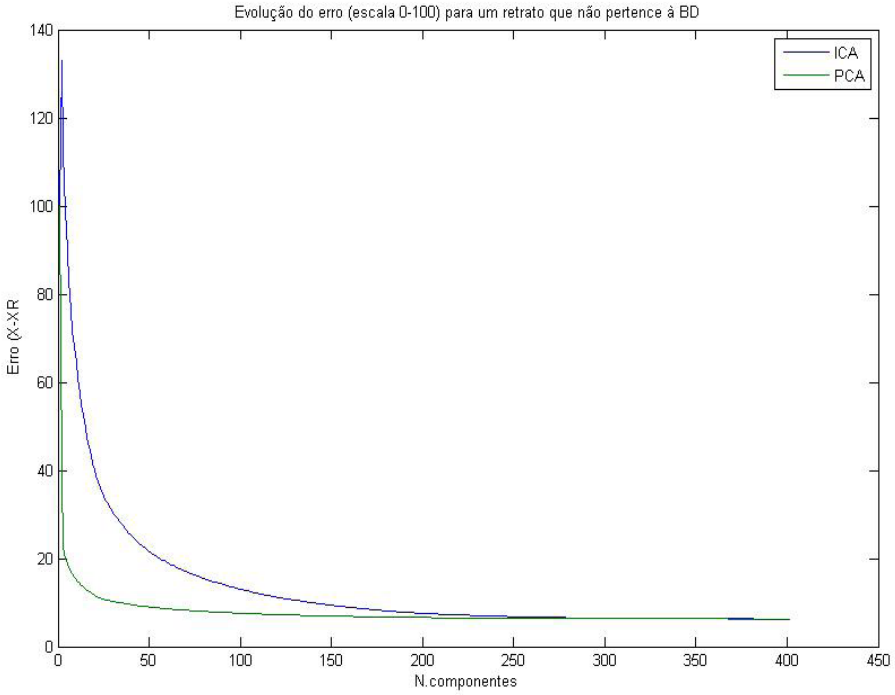


Figura 61 – Evolução do erro na reconstrução ICA e PCA para uma nova imagem (K=1:400)

A Figura 61 e a Figura 62 apresentam os resultados obtidos para um retrato que não pertence à BD, usando os eixos ICA e o algoritmo que apresentámos. A Figura 61 compara o erro de reconstrução com os resultados que obtivéramos para o PCA. O erro, como se observa, evolui de novo de forma idêntica nos dois casos, mas mais uma vez no PCA a convergência é mais rápida. Observa-se um valor idêntico de cerca de 6 para o erro final quer no PCA quer no ICA. A Figura 62 mostra alguns retratos reconstruídos com ICA e PCA com o mesmo número de componentes. Verifica-se que o processo de reconstrução, quase nada evolui a partir de 50 ou 100 componentes, respectivamente para o PCA e ICA (o que está em linha com a evolução do erro, que quase estabiliza para N entre 50 e 100 e entre 100 e 200, para cada um dos métodos).

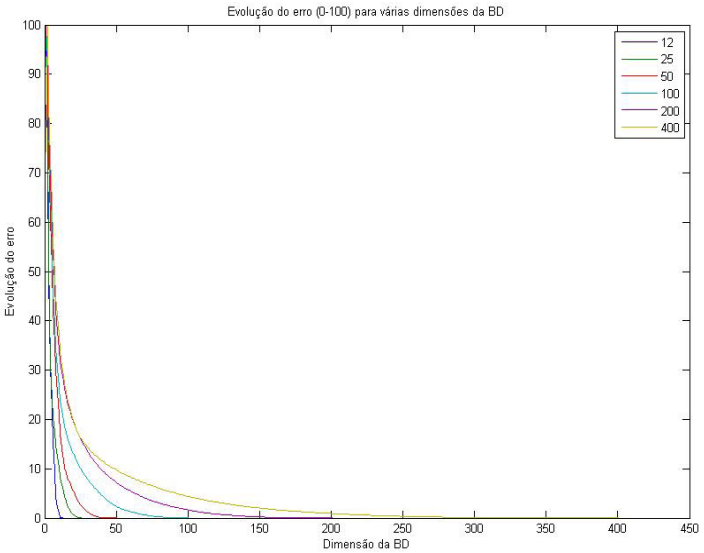
Figura 62 – Reconstrução de um retrato com K=20, 60,100 e 400 para PCA e ICA

Como se observa o retrato reconstruído com todos os componentes não só tem o mesmo erro (6 em 100), mas é idêntico em qualidade, para o PCA e para o ICA.

4.3.4 Dimensão da BD para o ICA

Os resultados obtidos com o PCA despertaram a nossa curiosidade para ver o que acontecia com o ICA quando se varia a dimensão da BD. Procurámos então analisar o comportamento do erro e da reconstrução de retratos, para BDs com dimensões diferentes, tal como fizéramos com o PCA, usando o ICA com o algoritmo *gauss*. As duas figuras a seguir mostram:

- a evolução do erro para BDs construídas com 400, 200, 100, 50, 25 e 12 retratos
- o nº de componentes necessários para reconstruir um retrato com erro < 1% para cada uma dessas BDs



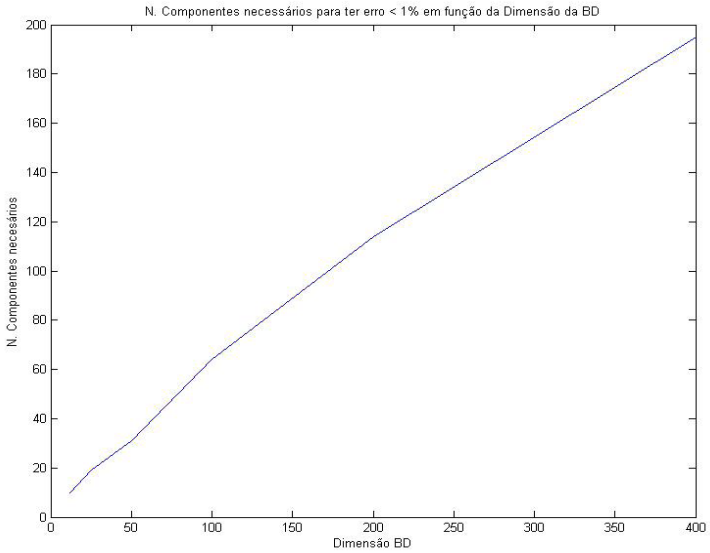


Figura 63 a b – Resultados obtidos com ICA para BDs de dimensão diferentes

Os resultados evidenciam que:

- tal como no PCA, o erro diminui mais rapidamente quando a dimensão da BD baixa
- aumentando a dimensão da BD as curvas coincidem cada vez mais até um determinado valor, por exemplo para 200 e 400 as curvas quase coincidem até um erro um pouco maior que 10%
- o nº. de componentes necessários para garantir um valor para o erro varia linearmente com a dimensão, tal como observáramos no PCA

Estes resultados podem mais uma vez surpreender mas demonstram apenas a capacidade do ICA, como método, tal como o PCA, para reconstruir as imagens sem erro, a partir dos componentes ICA que calcula.

Mas será interessante testar, tal como fizemos no PCA, a forma como a dimensão da BD afecta a reconstrução de retratos novos e comparar o valor final para o erro. Neste caso, em virtude de termos sistemas ICA, não ortogonais, tivemos de recorrer uma vez mais ao algoritmo *pesqproj*, para calcular os coeficientes de factorização do retrato novo utilizado nos testes, para cada dimensão da BD. A Figura 64 mostra a evolução do erro final de reconstrução de um retrato novo com a dimensão da BD. Verificamos que a curva do ICA quase coincide com curva para o PCA. Se tivermos em conta uma vez mais os resultados da secção 6, onde demonstramos que a zona crítica para o reconhecimento parece se situar para um erro > 9, poderemos concluir que neste caso (retratos com 200x200 pixéis) necessitamos, tal como no PCA, de um número de imagens na BD de cerca de 170.

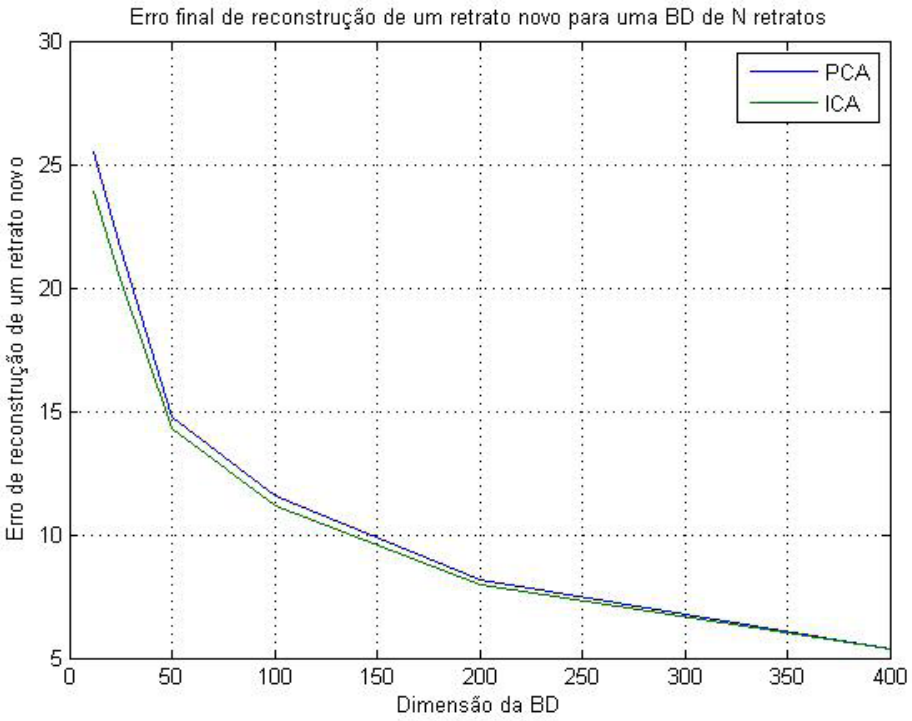


Figura 64 – Erro final de reconstrução de um retrato novo para uma BD de N retratos

Este limiar parece pois ser independente da técnica. Vai ser o limiar da dimensão da BD com retratos de 200x200 pixéis, para conseguir reconstruir retratos novos reconhecíveis.

(Hyvärinen and Oja 2000); (Hyvärinen, Oja, and Juha 2001)

4.4 Resultados com o PCA + ICA

Uma das vantagens do algoritmo *pesqproj* desenvolvido, descrito na secção anterior é a de possibilitar construir sistemas híbridos. Decidimos experimentar os resultados para um sistema com 800 eixos, associando os 400 componentes PCA com os 400 componentes ICA (*gauss*).

Observa-se que, de facto, para o retrato testado apenas 630 componentes, das 800, são utilizadas, após 1000 iterações do algoritmo *pesqproj*. A Figura 65 apresenta a evolução do erro de reconstrução que obtivemos, com o número de iterações do algoritmo *pesqproj*.

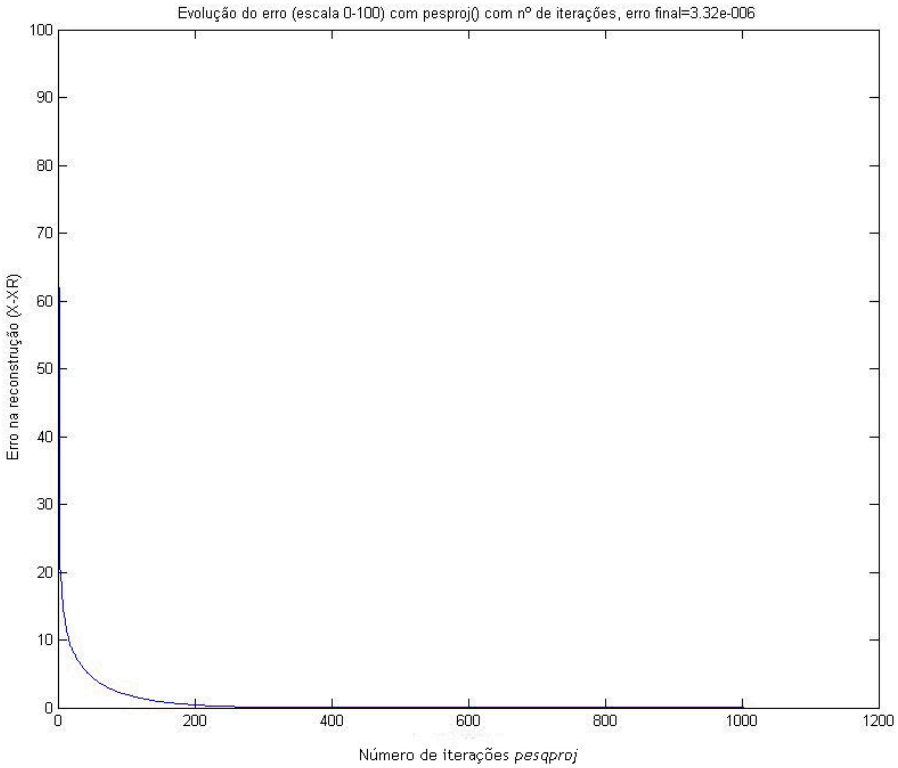


Figura 65 – Evolução do erro com o n. de iterações

Utilizámos outro script, idêntico ao que utilizámos na reconstrução PCA ou ICA, para fazer a reconstrução iterativa do retrato com os 630 componentes mantidos e utilizando os valores de coeficientes calculados pelo algoritmo *pesqproj*. Para isso, tal como se fez no PCA ou no ICA, utilizámos um vector de coeficientes A_K, obtido a partir do vector dos coeficientes para esse retrato A, calculados pelo *pesqproj*. A_K é facilmente obtido a partir de A, em cada iteração, fazendo zero todos os coeficientes de ordem> K, garantindo assim mais uma vez que na equação (49) que os componentes de ordem> K não são utilizados na reconstrução. É calculado o erro da reconstrução comparando X_R com X em cada iteração, utilizando a equação (48); cada vector retrato X_R assim reconstruído é recalculado como uma matriz imagem de 200x200 pixéis e guardado numa directoria de reconstruções com o formato tiff ou jpg.

A Figura 66 mostra a evolução do erro de reconstrução com o número de componentes utilizados para o sistema híbrido PCA+ICA e permite compará-lo com o erro no PCA.

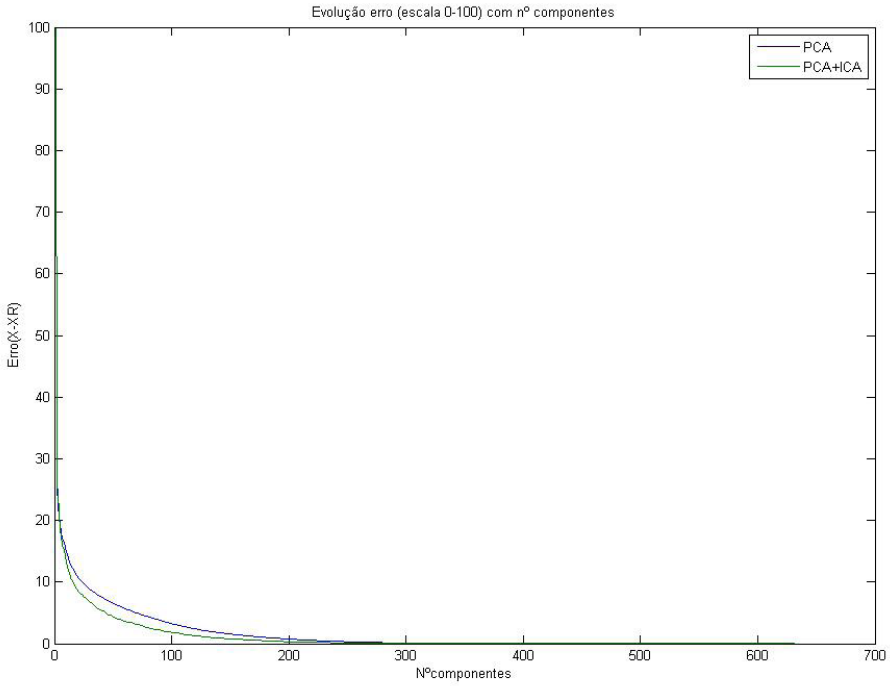


Figura 66 – Evolução do erro com o nº de componentes para PCA e ICA+PCA

Observa-se que o erro no sistema híbrido converge para zero mais rapidamente do que no PCA. A Figura 67 representa a diferença entre os dois erros. A menos de uma pequena flutuação logo no início, a diferença é sempre positiva, o que significa que se aumentou o carácter esparso do processo, com a utilização PCA+ICA, consegue-se o mesmo erro com menos componentes.

Este resultado era expectável pois no algoritmo, ao seleccionar por ordem decrescente as projecções mais significativas, confirmámos que dava preferência muitas vezes a componentes ICA, em vez de ficar agarrado à ortogonalidade do PCA. Para a imagem seleccionada a ordem dos primeiros 20 componentes seleccionados pelo *pesqproj* é:

1 2 6 403 477 407 7 8 33 442 429 14 425 9 20 424 15 445 57 447[11]

Ou seja, observa-se que os 3 primeiros componentes são PCA, mas os 3 seguintes são componentes ICA, e que nos vinte primeiros componentes, há 11 PCAs e 9 ICAs.

11 Poderá surpreender aparecer uma ordem 1 2 6 para os componentes PCA, em vez de 1 2 3, que seria ordem dos valores próprios, mas é preciso ter em conta que o algoritmo encontra a ordem mais favorável para um retrato e os valores próprios correspondem à ordem mais favorável para o conjunto de todos os retratos.

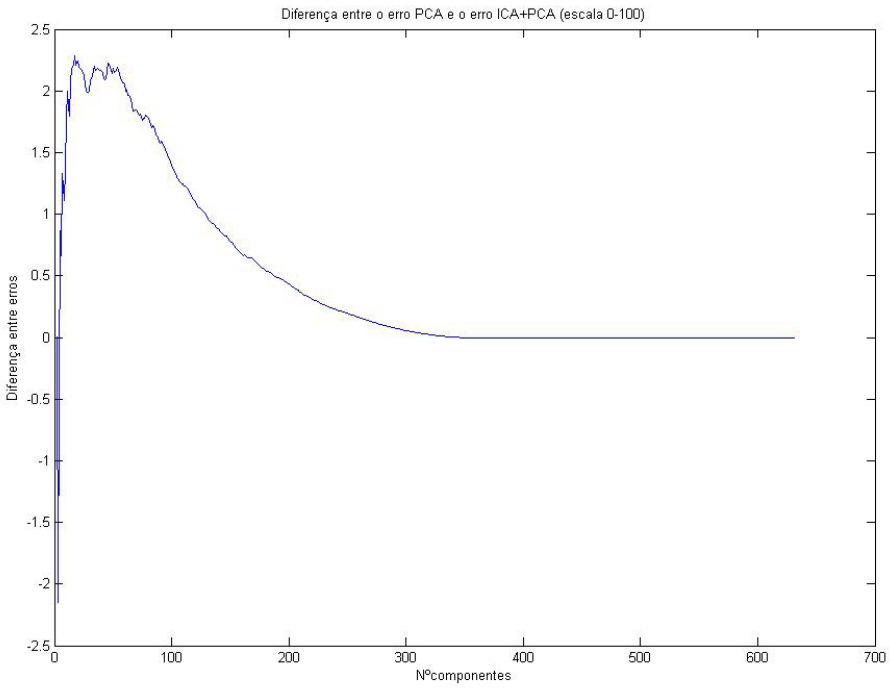


Figura 67 – Diferença entre o erro para o PCA e o erro para PCA+ICA

A melhoria introduzida pelo sistema híbrido PCA+ICA permite de alguma forma suportar a ideia de complementaridade dos dois métodos.

4.5 Resultados com o NMF

Desenvolveram-se um conjunto de scripts em MATLAB, recorrendo a vários algoritmos NMF, que utilizam métodos distintos: métodos multiplicativos, de mínimos quadrados alternados e de gradientes projectados, para fazer a decomposição da matriz de vectores de imagens em componentes NMF. As secções seguintes apresentam e comparam os resultados que obtivemos com os diferentes métodos. Os scripts Matlab de cálculo dos componentes NMF para uma matriz de retratos X e de reconstrução iterativa de um retrato com K componentes têm uma lógica muito parecida com os que descrevemos para o PCA e para o ICA, a menos das particularidades do NMF. Por exemplo é necessário no NMF inicializar as matrizes W e H antes de utilizar os algoritmos NMF e não se pode retirar o valor médio a cada vector X nem ao conjunto dos dados, sob pena de termos valores negativos, que o NMF proíbe.

4.5.1 Método dos gradientes projectados

Utilizámos o algoritmo dos gradientes projectados (GP) de Chih-Jen Lin. Para utilizar este algoritmo precisamos de inicializar as matrizes W e H e o algoritmo utiliza o método dos gradientes projectados para minimizar a função f(W,H) e calcular iterativamente as matrizes W e H.

Observámos que o algoritmo parece convergir sempre, ou seja, parte de um valor para a norma do gradiente e obtém no fim um valor bastante mais baixo. No entanto observámos também que os resultados dependem muito da inicialização feita para as matrizes W e H, podendo mesmo obter-se resultados sem grande sentido, por exemplo todos os componentes NMF correspondendo a uma imagem cinzenta, sem informação, ou todos parecidos com uma imagem próxima da média das

imagens. Outro parâmetro difícil de controlar como veremos é o número de componentes NMF a calcular.

Conseguem-se, para algumas inicializações de W e de H, obter componentes imagéticos como os representados na Figura 68, mas mesmo neste caso o erro da reconstrução varia bastante, dependendo dos valores utilizados para a inicialização das matrizes. Não conseguimos, mesmo junto do autor do algoritmo, qualquer explicação para este fenómeno, nem conseguimos verificar com rigor qual a inicialização mais indicada.

A Figura 68 apresenta os componentes NMF – no caso de calcularmos 200 componentes – e a Figura 69 apresenta a evolução da imagem reconstruída com K componentes, K=1:200. Correspondem aos melhores resultados que obtivemos para o erro, depois de várias tentativas.

Figura 68 – Componentes NMF – método GP

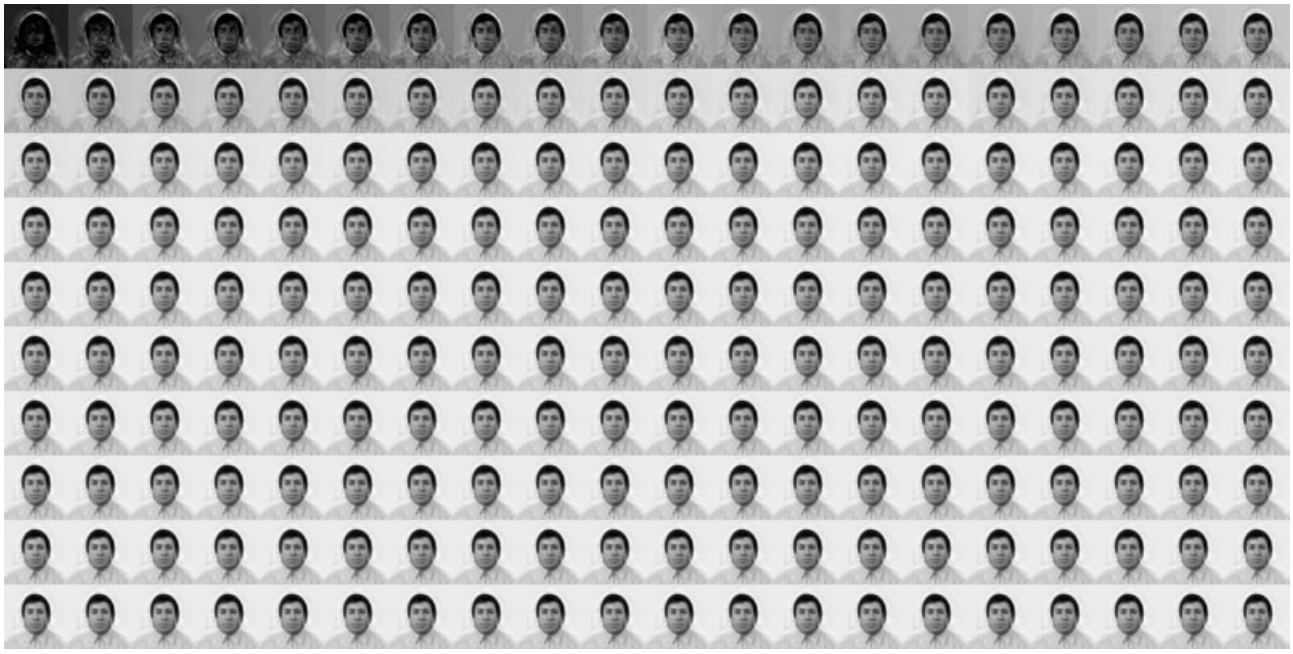


Figura 69 – Imagem recuperada com o NMF

Com este método obtivemos um erro final significativo, como se observa na Figura 70, de cerca de 8 em 100, neste caso para 100 componentes e 1000 iterações. Verifica-se também que esse erro se consegue com pouco mais de 30 componentes. Em rigor, para o retrato que utilizámos neste teste, apenas 47 dos coeficientes Hi são maiores do que zero, como se observa na Figura 71, o que traduz o carácter esparso do processo.

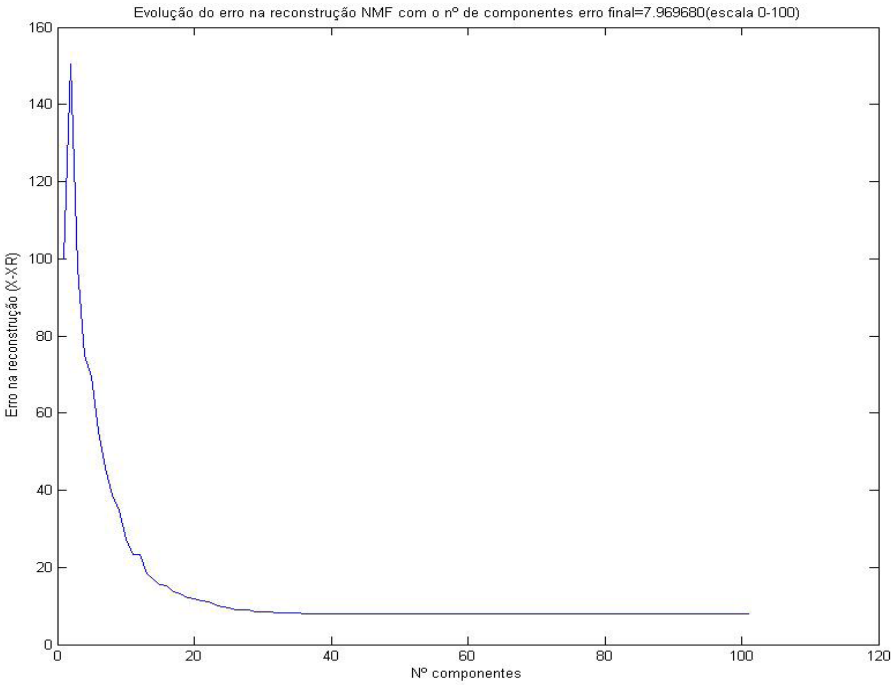


Figura 70 – Erro na reconstrução para o NMF com método dos gradientes

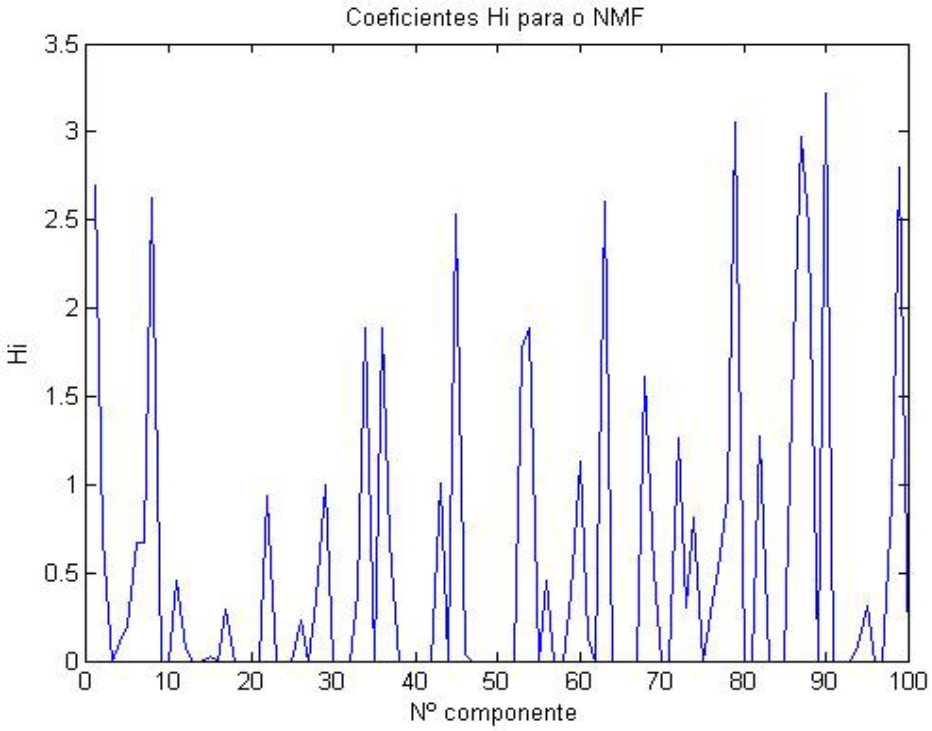


Figura 71 – Coeficientes Hi para o o NMF – método GP

Estes primeiros resultados permitem-nos concluir que o NMF que utiliza gradientes projectados possibilita obter matrizes W e H que aproximam a matriz X, ainda que com um erro final bastante

grande, que dificulta mesmo o reconhecimento de um ponto de vista fotográfico.

O algoritmo é relativamente rápido e permite resultados razoáveis com poucas iterações.

O erro final depende muito da inicialização e parece ser difícil de controlar. Não podemos portanto garantir que não haja outras inicializações que conduzissem a um erro menor do que o que obtivemos.

Os componentes que se encontram parecem ter características globais – são retratos com alguma analogia com as Eigenfaces do PCA – e não confirmam a tese de que o NMF permitiria obter componentes orientados a zonas do rosto, o que os autores que defendem o NMF para reconhecimento de rostos apresentam como uma das suas grandes vantagens.

(Chih-Jen 2007)

4.5.2 Método multiplicativo

Os resultados foram obtidos com o algoritmo disponibilizado pela Universidade Técnica da Dinamarca, suportado no método multiplicativo de Lee e Seung. Observámos que o algoritmo também parece convergir sempre, ou seja, o valor do erro da aproximação NMF parece tender para zero com as iterações, ainda que o processo seja lento, pois exige um elevado número de iterações. O algoritmo é muito mais lento do que o que utiliza o método dos gradientes: para a dimensão da nossa BD e com mil iterações o algoritmo demora horas a terminar, em contraste com os poucos minutos do método dos gradientes. Observa-se por outro lado que o resultado parece não depender da inicialização das matrizes W e H, tal como acontecia no caso dos gradientes. Experimentámos várias inicializações, inclusive algumas que não funcionaram com os gradientes e o resultado final que obtivemos foi sempre parecido: o valor final do erro da aproximação de X com as matrizes W e H, calculado pelo algoritmo MM, usando mil iterações foi sempre de cerca de 1.6 (distância euclidiana), e o erro final da reconstrução do retrato, calculado por nós, com base na equação (48) foi sempre de cerca de 4 em 100. A Figura 72 representa a evolução do erro final na reconstrução de um retrato, utilizando o algoritmo avaliado, com 1000 iterações e 400 componentes NMF, para três inicializações diferentes de W e H, inicializações que no método dos gradientes conduziram a resultados bastante diferentes e mesmo a uma situação onde não foi possível fazer a reconstrução. Para cada uma das inicializações utilizou-se o algoritmo MM para calcular as matrizes W e H e depois utilizámos um script Matlab, idêntico ao que descrevemos para PCA ou ICA, de reconstrução iterativa do retrato com K componentes, com K=1:400. Observa-se que as curvas do erro de reconstrução de um retrato diminuem de uma forma quase monótona com o nº de componentes e que não diferem muito para os 3 casos, convergindo, quase coincidindo a partir de K=200.

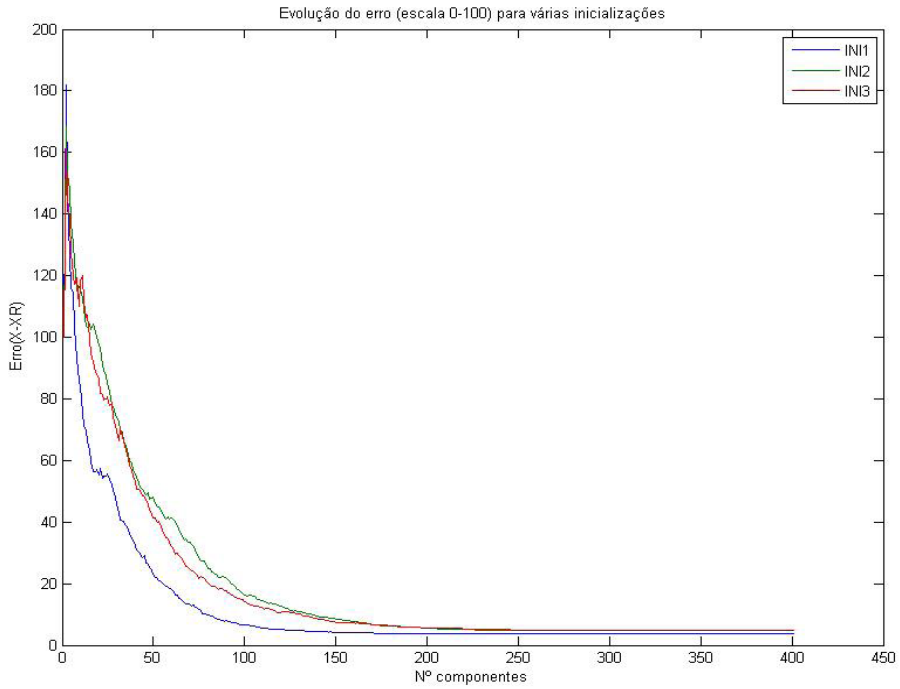


Figura 72 – Erro na reconstrução NMF – método multiplicativo com várias inicializações

Verifica-se ainda que o erro final, na aproximação de X por W*H pelo método multiplicativo, calculado pelo algoritmo, diminui de uma forma monótona com o número de iterações do algoritmo, e que o erro final de reconstrução de um retrato, calculado por nós, com base na equação (48), também diminui, mas a diminuição deste último torna-se muito lenta a partir de um determinado valor do número de iterações. Por exemplo, experimentámos fazer 1500 iterações no cálculo de W e de H, em vez de 1000, e a distância euclidiana calculada pelo algoritmo do MM, entre X e W*H, baixou de 1.67 para 1.49, mas o erro final na reconstrução de um retrato baixou apenas de 4.87 para 4.34. Experimentámos por fim 5000 iterações e o algoritmo do MM calculou uma distância euclidiana entre X e W*H de 1.12 e o nosso script obteve erro final na reconstrução do mesmo retrato de 3.84. Qual o erro que se conseguiria aumentando indefinidamente as iterações? Zero? Não conseguimos responder qual é exactamente o valor para que o erro final tende, aumentando o número de iterações.

Os componentes imagéticos NMF obtidos pelo método multiplicativo são idênticos aos que se obtiveram com o método dos gradientes, ainda que pareça aqui existir alguma localização desses componentes em zonas da face (ver manchas brancas na Figura 73). A reconstrução iterativa de um retrato é em tudo idêntica à que observámos no método dos gradientes como se vê na Figura 74.

Figura 73 – Componentes NMF com o método multiplicativo

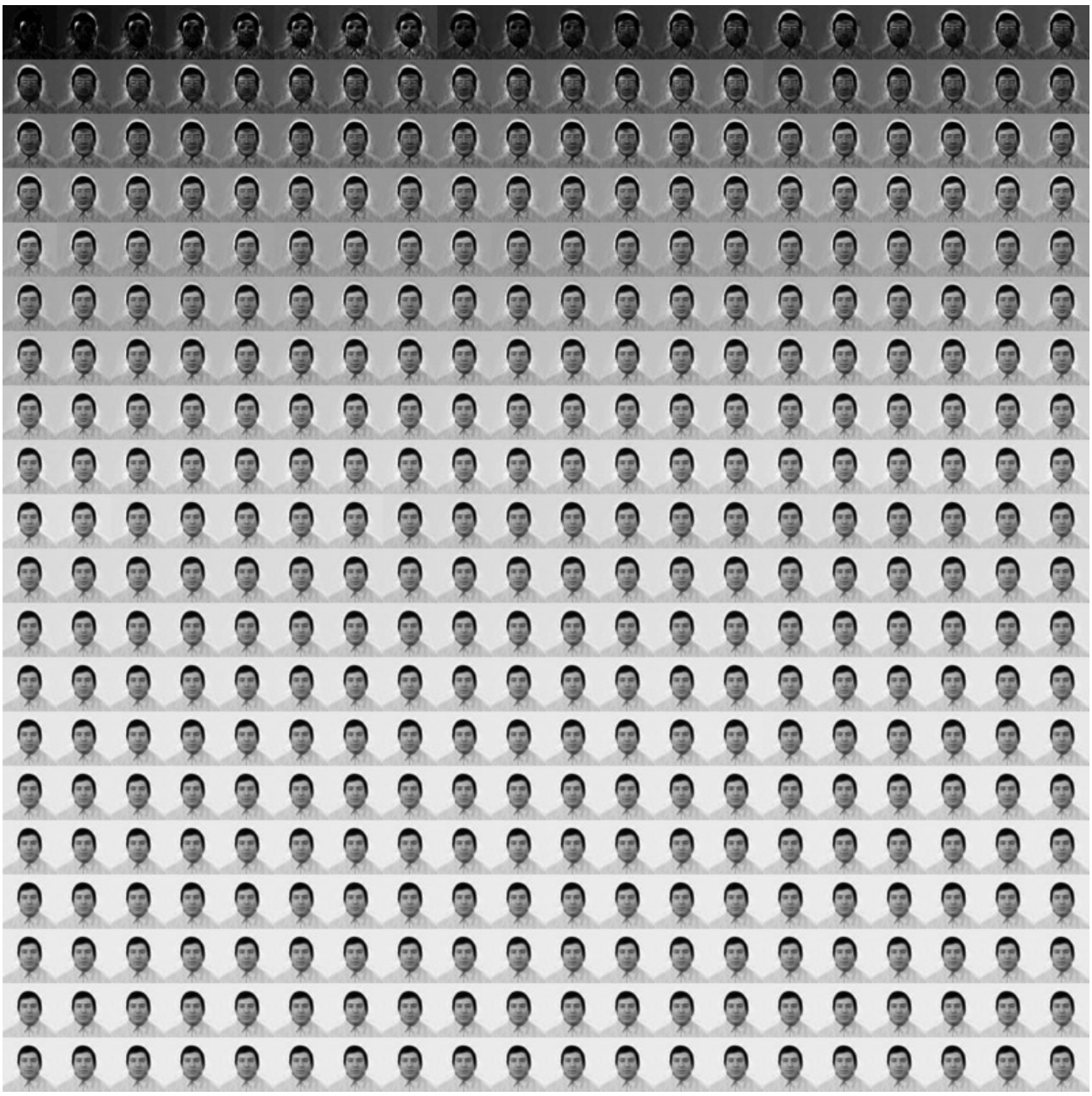


Figura 74 – Reconstrução de um retrato com o método multiplicativo

Observa-se no entanto uma menor ‘esparsidade’ do que a que observámos no método dos gradientes. O número de H_i iguais a zero é muito menor, como se observa na Figura 75.

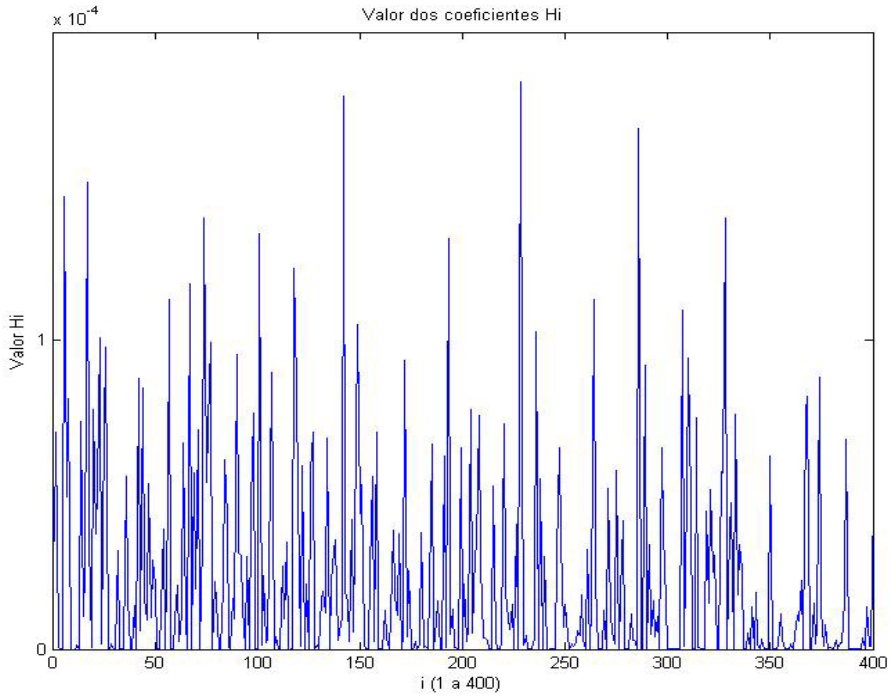


Figura 75 – Coeficientes Hi para o o NMF com método multiplicativo

Para o retrato que utilizámos neste teste, com 1000 iterações, precisamos de mais de 130 componentes em 400 para obter um erro menor do que 10.

Estes primeiros resultados permitem-nos concluir que o NMF baseado no método multiplicativo também possibilita obter matrizes W e H que aproximam a matriz X, ainda que também com um erro final diferente de zero, implicando uma reconstrução imperfeita, mesmo com 5000 iterações, o que correspondeu a mais de doze horas de processamento. No entanto este erro diminui ligeiramente com número de iterações, ainda que muito lentamente a partir de um determinado valor, e não somos capazes de dizer qual o valor final do erro se o nº de iterações aumentasse indefinidamente.

O algoritmo parece ser imune à inicialização de W e H, desde que esta cumpra a restrição de positividade. Obtivemos resultados idênticos para inicializações muito diferentes.

O algoritmo é bastante lento, se comparado com o que utiliza gradientes projectados.

(Lee and Seung 2001);

4.5.3 Método dos mínimos quadrados alternados

Os resultados obtidos com este método para a nossa BD de retratos são no mínimo estranhos. Observa-se que as três grandezas, que são calculadas pelo próprio algoritmo, para controlar o erro da iteração

- diff = sum(sum(abs($X_{R_anterior}$-X_R))); onde X_R é a matriz X reconstruída = W*H
- distância euclidiana entre X e X_R
- error=mean(mean(abs(X-W*H)))/mean(mean(X));

variam de uma forma inconstante ao longo do processo de iteração, ora diminuindo ora aumentando.

Os componentes, a olho nu, que calculámos com o nosso script, são no entanto parecidos com os obtidos por outros métodos, como se observa na Figura 76.

Figura 76 – Componentes NMF com o método mínimos quadrados alternados

Observamos no entanto que há componentes nulos e observando mais de perto verificamos que cada componente parece ser uma espécie de esboço modernista de cada elemento da BD. Ou seja, este método parece conduzir à solução trivial em que cada componente NMF é algo próximo de cada retrato.

A curva para o erro, algo estranha, obtida para 400 componentes, levou-nos a investigar o que aconteceria com menos componentes. A Figura 77 representa a evolução do erro final para a reconstrução do mesmo retrato, com 20, 30, 50, 100 e 400 componentes.

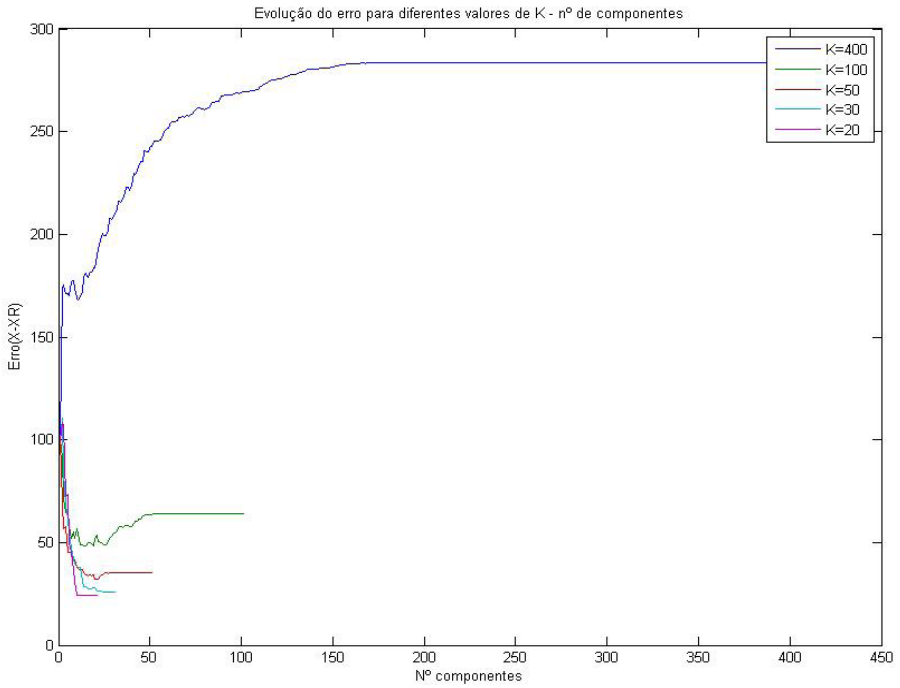


Figura 77 – Erro na reconstrução NMF com método mínimos quadrados alternados

Com 400 componentes observa-se que o erro final é enorme e a curva tem um comportamento anormal e o erro aumenta em vez de diminuir. O resultado é melhor se diminuirmos o nº de componentes calculados, o que também é estranho. Com 20, 30, 50 componentes a curva é quase monótona decrescente, mas mesmo nestes casos o erro é muito grande, acima de 20 em 100.

A reconstrução é para este método desastrosa, como se observa na Figura 78.

Figura 78 – Retrato reconstruído – método dos mínimos quadrados -18 e 400 componentes

O algoritmo é também bastante lento se comparado com o que utiliza gradientes projectados.

Estes primeiros resultados permitem-nos concluir que este método parece não permitir reconstruir retratos com erros minimamente aceitáveis.

4.5.4 Comparação entre os diferentes métodos NMF

Resumimos nesta secção as conclusões que pudemos tirar da avaliação que fizemos para a nossa BD do NMF e dos vários métodos propostos, avaliando comparativamente a forma como permitem decompor e reconstruir retratos da BD. Utilizámos o algoritmo dos gradientes projectados (GP) de

Chih-Jen Lin e os algoritmos propostos pela Universidade Técnica da Dinamarca que implementam o método multiplicativo (MM) de Lee e Seung e o método dos mínimos quadrados alternados (MQA). As conclusões estão organizadas em vários itens:

- aspecto dos componentes
- variação dos resultados com o número de iterações
- sensibilidade às inicializações
- rapidez do algoritmo
- erro de reconstrução e resultados da reconstrução de um retrato

4.5.4.1 Aspectos dos componentes no NMF

Todos os métodos conduzem a componentes que lembram caras algo fantasmagóricas (*NMF-faces)*, mas com um aspecto muito diferente das eigenfaces do PCA. Tal como se observa na Figura 79 os componentes NMF para os três métodos parecem ser do mesmo tipo numa primeira observação, mas o método MM parece ser aquele onde se consegue identificar uma localização em zonas específicas da face, para cada um dos componentes.

Figura 79 – Componente nº 99 para cada um dos métodos NMF: GP, MM e MQA

4.5.4.2 Erro e resultados na reconstrução[12] no NMF

O método MQA conduz a erros muito elevados e não permite sequer a reconstrução. Por outro lado o erro aumenta com o nº de componentes. O método multiplicativo parece ser aquele que permite erros ligeiramente mais baixos (4 em 100 para 1000 iterações, contra 6 do GP – ver tabelas abaixo). Esta conclusão não pode ser no entanto definitiva, devido à dependência das inicializações e do número de iterações. Mas o método GP permite um decrescimento mais rápido do erro, como se observa na Figura 80.

12 O facto de o NMF utilizar matrizes positivas, que integram as médias, obrigou-nos a ter algum cuidado no cálculo do erro, para garantir coerência com os cálculos do erro para o PCA e ICA. O erro é sempre calculado como a energia contida no vector erro=$X-X_R$, onde os dois vectores, X e X_R, são normalizados a 1. No NMF estes dois vectores têm de ser convertidos nos vectores X e X_R sem média, para poder comparar o valor do erro no NMF com os valores calculados em PCA e ICA.

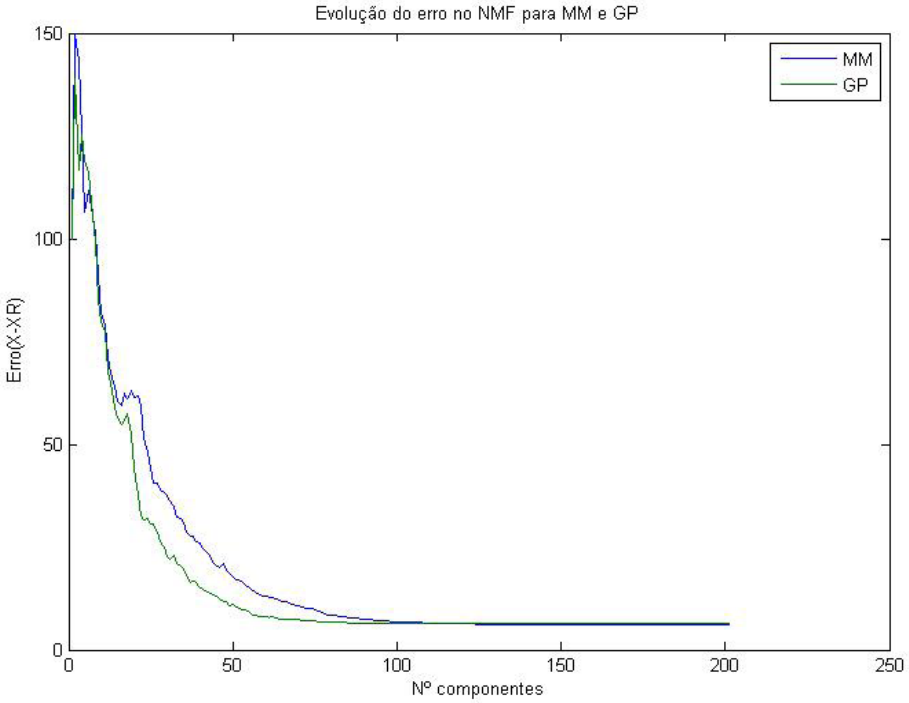


Figura 80 – Evolução erro em MM e GP para nIter=1000

O erro é no entanto sempre significativo, para qualquer um dos métodos, pelo menos para as iterações que experimentámos, o que não permite nem de longe reconstruções perfeitas. O NMF parece não ser portanto indicado para o reconhecimento e reconstrução de rostos, quando comparado ao PCA ou ao ICA. Iremos analisar se permite melhorar os resultados, se associado ao PCA e ao ICA num sistema híbrido.

4.5.4.3 Número de iterações no NMF

A Tabela 3 e a Tabela 4 resumem os resultados obtidos em função do número de iterações. A Tabela 3 dá-nos o erro final conseguido, a Tabela 4 o número de componentes necessários, para o erro não ser superior ao valor imposto.

Método	**Condições**	**Iter=100**	**Iter=1000**	**Iter=1500**	**Iter=5000**	**Iter=10000**
GP(errofinal)	N=400 K=200	6.10	6.43			5.07
MM(errofinal)	N=400 K=400		4.87	4.34	3.84	
MQA(errofinal)	N=400 K=400		>100			

Tabela 3 – Variação do erro no NMF com o nº de iterações
N – dimensão da BD
K – nº de componentes

Observa-se na Tabela 3 que para o método GP o erro parece não depender muito do nº de iterações, variando em torno de 6. A Figura 81 mostra claramente que as três curvas para nIter=100, 1000, 10000 não diferem muito, conseguindo-se apenas um decrescimento mais rápido para os valores mais altos do nº de iterações mas para 100 e 1000 iterações a diferença é muito pequena.

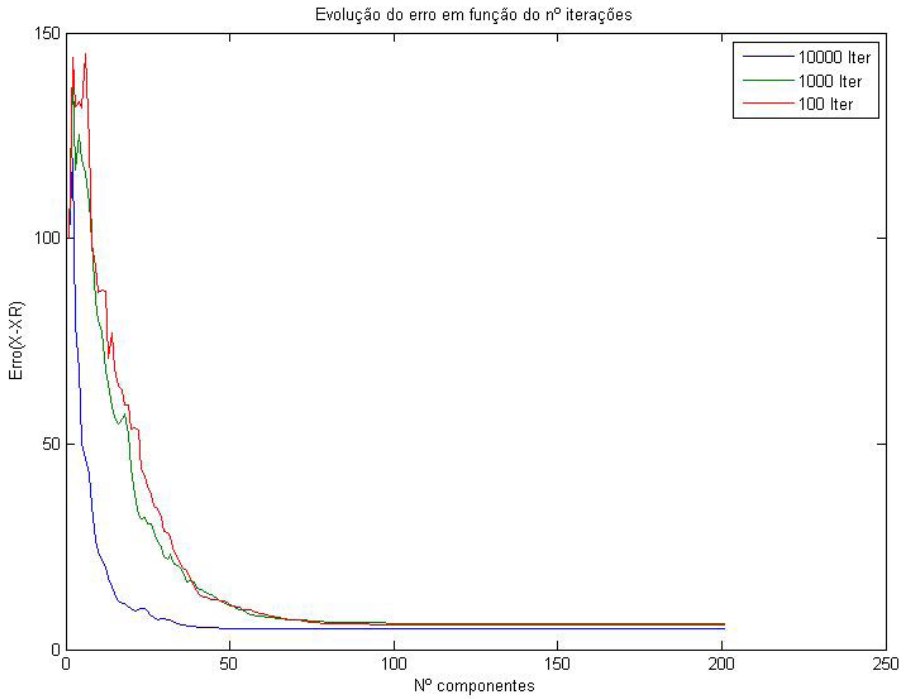


Figura 81 – Evolução erro em GP para nIter=100, 1000,10000

Para o MM o erro final parece decrescer lentamente com o nº de iterações (ver Tabela 3), mas um grande nº de iterações introduz um peso muito grande no processamento. Tal como no GP, aumentando o nº de iterações, o decrescimento torna-se um pouco mais rápido, como se observa na Figura 82.

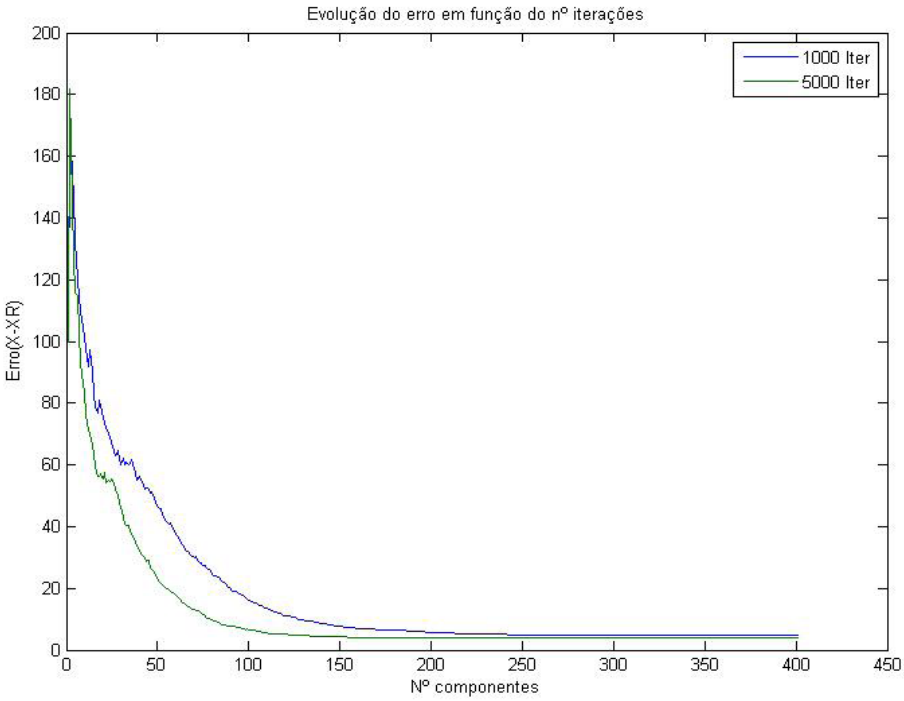


Figura 82 – Evolução erro em MM para nIter=1000, 5000

Quanto ao nº de componentes necessários para ter um erro <10 em 100, confirma-se na Tabela 4 que no método GP o número de iterações parece não introduzir uma variação significativa na forma como a curva converge, para 100 e 1000 iterações, tal como já observáramos na Figura 81,

pois as duas curvas (Iter=100,1000) quase coincidem. Mas em MM aumentando o nº de iterações, para além de baixar o erro final, a curva torna-se mais rápida a convergir, tal como se observa na Figura 82 e na Tabela 4.

Método	Condições	Iter=100	Iter=1000	Iter=1500	Iter=5000	Iter=10000
GP(NComp)	N=400 K=200 Tol=0.00000001 Erro<10	53	52			19
MM(NComp)	N=400 K=400 Erro<10		127		78	

Tabela 4 – Nº componentes NMF para assegurar um erro < 1 em função do nº iterações
N – dimensão da BD
K – nº de componentes

Pode-se colocar a questão de qual o número de iterações a utilizar, quer num quer noutro método: tentar definir nos métodos MM e GP qual o nº de iterações que se justifica fazer, para optimizar o erro final e tornar a construção do retrato mais rápida, com o preço a pagar no tempo de processamento. A decisão não parece ser fácil.

4.5.4.4 Número de componentes no NMF

A Figura 83 mostra a evolução do erro para GP e MM, variando o número de componentes.

As Tabela 5 e a Tabela 6 resumem os resultados para os 3 métodos, variando o nº de componentes calculados (K), ou seja a dimensão da matriz W, que é NxK.

Método	K=30	K=50	K=100	K=200	K=400
GP(errofinal)		9.78	7.97	6.43	-
MM(errofinal)			7.94	6.08	4.87
MQA(errofinal)	>20	>20	>50		>100

Tabela 5 – Variação do erro final com o nº de componentes NMF com 1000 iterações[13]

O método MQA tem um comportamento estranho e permite melhores resultados com 20, 30 ou 50 componentes do que com 100 ou 400. Parece ser um método sem interesse para nós. Por isso não fizemos com ele algumas das comparações. Nos métodos GP e MM o erro final melhora aumentando o nº de componentes, como se observa nas figuras, mas por outro lado, talvez surpreendentemente, aumentando o nº de componentes o erro diminui mais lentamente (ver a Tabela 6).

13 Não se calcularam 400 componentes em GP por falta de memória na máquina utilizada para correr o algoritmo

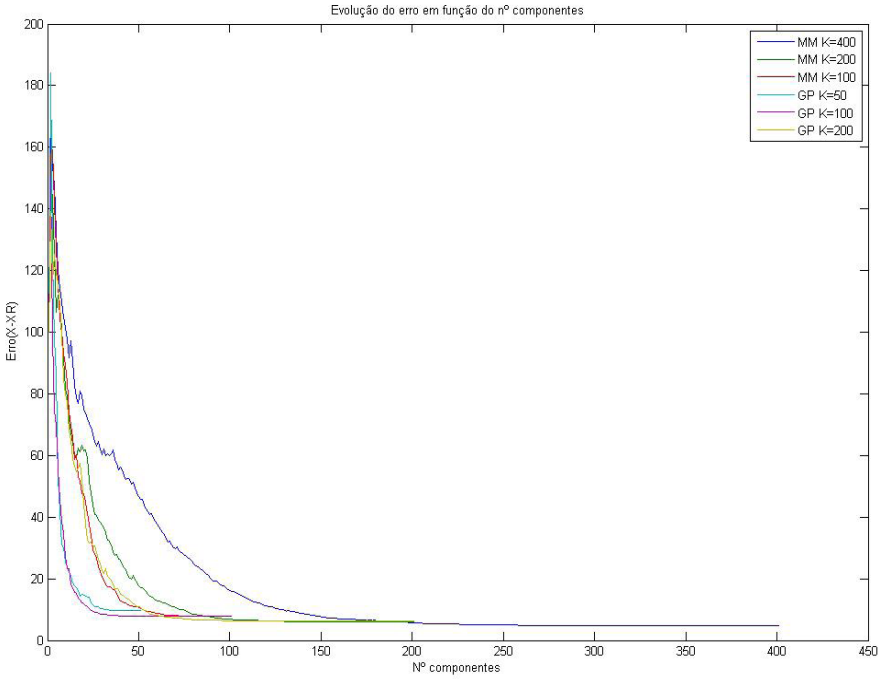


Figura 83 – Evolução do erro em GP e MM para K=50, 100, 200, 400

Método	K=400	K=200	K=100	K=50	K=30
GP(NComp) Erro<10	-	52	23	31	
MM(NComp) Erro<10	127	74	52		

Tabela 6 – Nº componentes NMF para assegurar um erro < 10 em função de K[14]

A diminuição da 'esparsidade' com K é clara em MM. Em GP o resultado para K=100 é o melhor de todos. Há aqui outro parâmetro de difícil escolha. Qual o K ideal para cada método? Qual o número de componentes a calcular? Aumentando K diminui o erro final mas pode também diminuir a 'esparsidade', ou não, no caso do GP. Verifica-se que o método GP conduz a erros ligeiramente maiores, mas a uma maior 'esparsidade' (são precisos menos componentes), tal como nós pretendemos.

4.5.4.5 Erro fotográfico no NMF

O erro geométrico pode ser medido utilizando múltiplas definições matemáticas, como vimos em 3.2.1, mas também nos interessa o erro fotográfico, que é um erro subjectivo, relacionado com a qualidade da reconstrução fotográfica dos diferentes processos. A Figura 84 representa a imagem reconstruída para o mesmo retrato, com 1000 iterações, para valores do erro de 10, e para o erro final para as diferentes técnicas utilizadas pelo NMF: método dos gradientes projectados, multiplicativo e dos mínimos quadrados alternados.

14 Não se calcularam 400 componentes em GP por falta de memória na máquina utilizada para correr o algoritmo

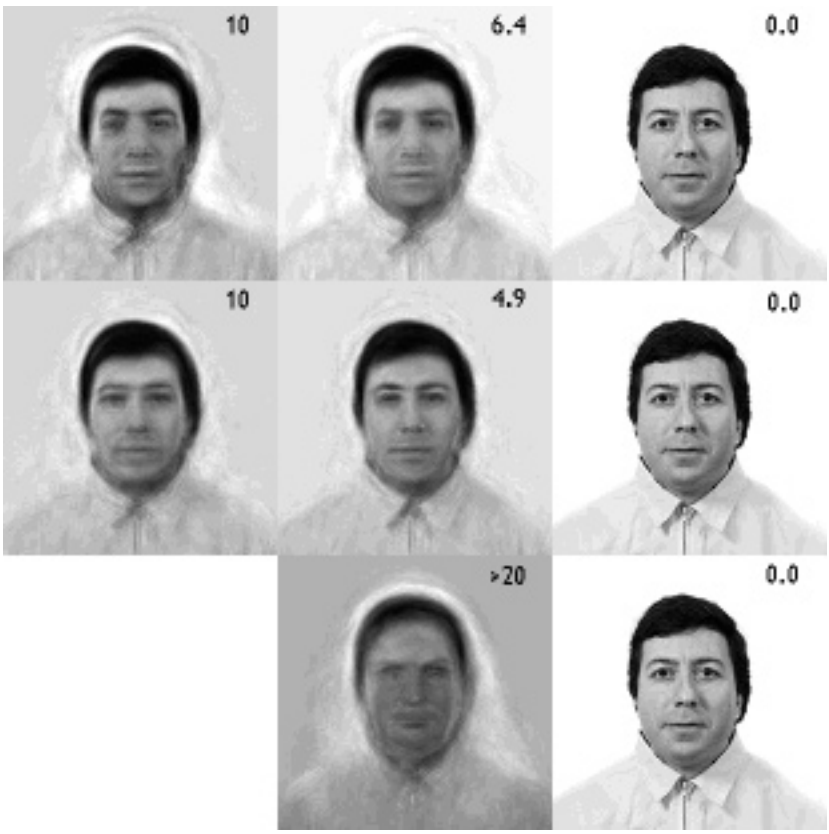


Figura 84 – Qualidade fotográfica para o NMF: GP, MM e MQA
GP com K=200
MM com K=400
MQA com K=50

No canto superior direito de cada retrato temos o erro na escala de 0 a 100. A coluna do meio apresenta a reconstrução final do NMF, com todos os componentes, para cada um dos três métodos testados. Os retratos da direita são originais.

De realçar a relativa falta de qualidade fotográfica da reconstrução de um retrato, se compararmos com o PCA ou com o ICA, apesar de parecer continuar a permitir um reconhecimento sem grandes dificuldades, se exceptuarmos o métodos dos mínimos quadrados que como vimos tem erros muito grandes. Observamos ainda que para os 2 métodos, GP e MM, o mesmo erro geométrico parece conduzir a uma qualidade fotográfica idêntica.

4.5.4.6 Inicializações no NMF

O método de gradientes demonstrou ter uma enorme dependência das inicializações, o que introduz outra variável de controlo muito difícil. Algumas inicializações parecem impedir mesmo a convergência. Pelo contrário o método multiplicativo demonstrou obter resultados idênticos após um certo número de iterações, independentemente das inicializações.

4.5.4.7 Rapidez no NMF

O método multiplicativo é claramente mais lento do que o método dos gradientes para o mesmo número de iterações.

4.5.4.8 Avaliação das estatísticas dos coeficientes no NMF

O método dos gradientes projectados parece conduzir a uma 'esparsidade' maior do que o método multiplicativo. A curva de erro converge mais rapidamente no método dos gradientes e portanto o número de componentes necessários para ter um erro próximo do erro final é menor. Para ajudar a analisar e a comparar os resultados obtidos para os coeficientes de mistura, calculados com cada um dos métodos utilizados, analisámos mais uma vez a dispersão desses coeficientes e as estatísticas de 3ª e de 4ª ordem, a estatística de assimetria (*skewness)* e de 'achatamento' (*kurtosis)*.

$$X=W*H \qquad (54)$$

A partir da expressão de cálculo de X em NMF – equação (54) – verificamos que cada retrato da BD é calculado com o vector de coeficientes H_{ij}, uma coluna da matriz H. A partir das matrizes H calculadas pelos algoritmos, para os dois métodos, GP e MM, ambos com 100 componentes e com 1000 iterações, obtivemos os histogramas para a distribuição dos coeficientes H_{ij} por coluna e para as respectivas variáveis estatísticas de assimetria e de 'achatamento'. Apresentamos na Figura 85 e na Figura 86 os 3 histogramas, respectivamente para o GP e MM. Ao analisar a estatística dos dados, em particular o valor de *kurtosis*, que mede de alguma forma a 'esparsidade' do processo relativa aos componentes, verificamos que o valor de *kurtosis* para o GP é de 366 enquanto no MM é apenas de 12.

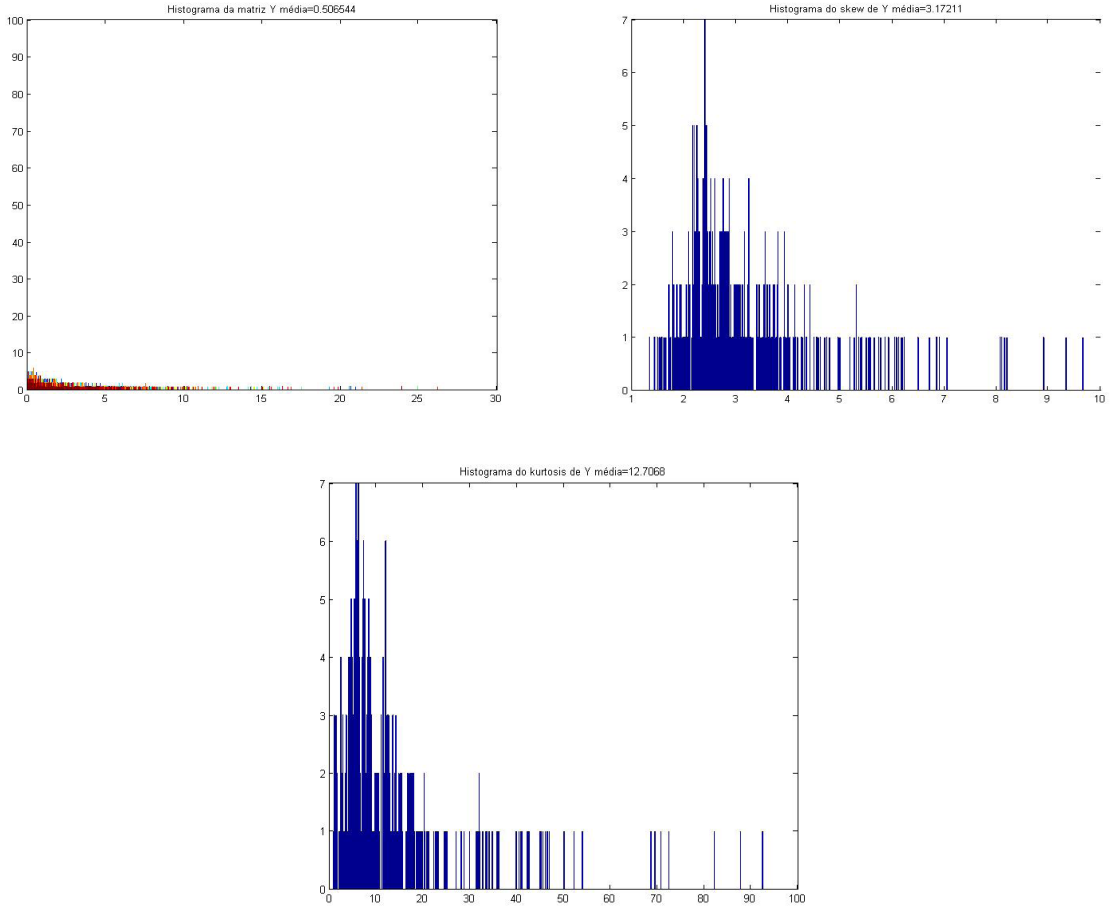


Figura 85 – Histogramas para o método GP
Coeficientes, skewness e kurtosis

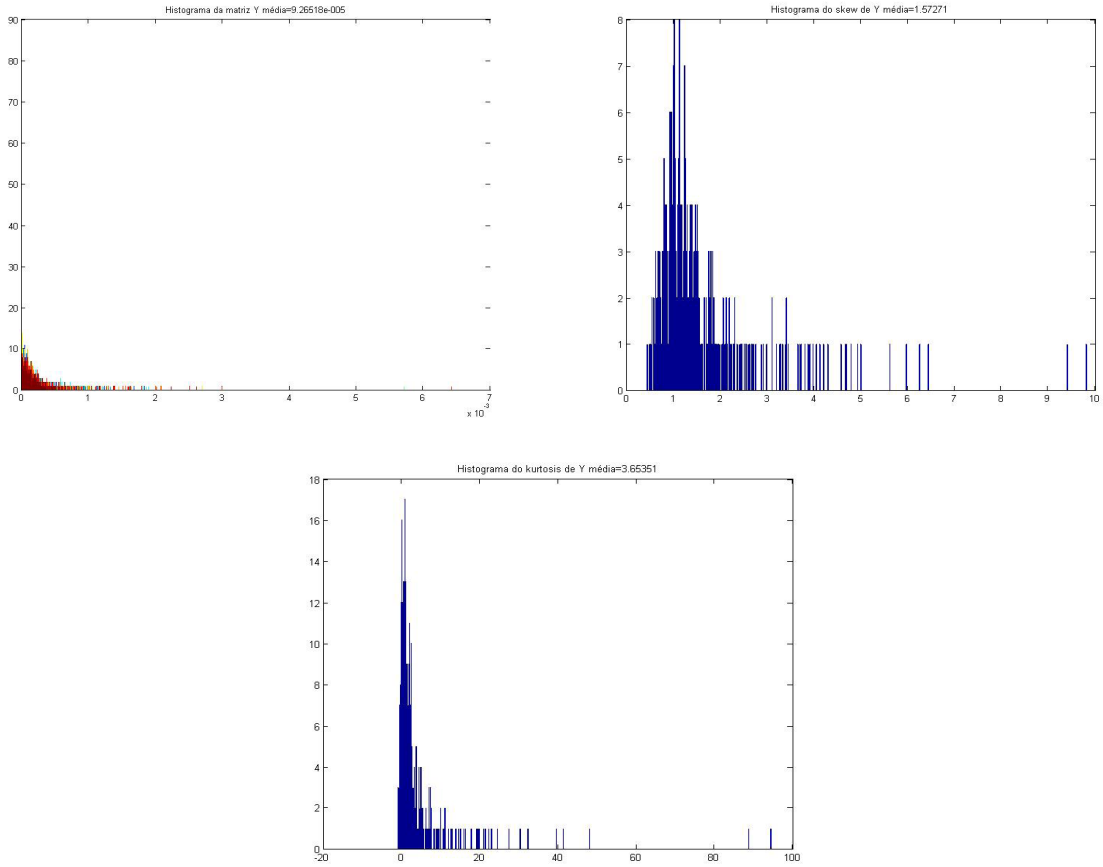


Figura 86 – Histogramas para o método MM
Coeficientes, skewness e kurtosis

No caso do método GP verifica-se ainda neste exemplo, que só cerca de 1/3 dos coeficientes H_{ij} (14669 em 40000) são diferentes de zero. Já no caso do método MM verifica-se que todos os coeficientes H_{ij} são maiores do que de zero, ainda que alguns muito pequenos. A diferença pode observar-se na Figura 87, que confirma os resultados: a 'esparsidade' do GP é maior do que a do MM.

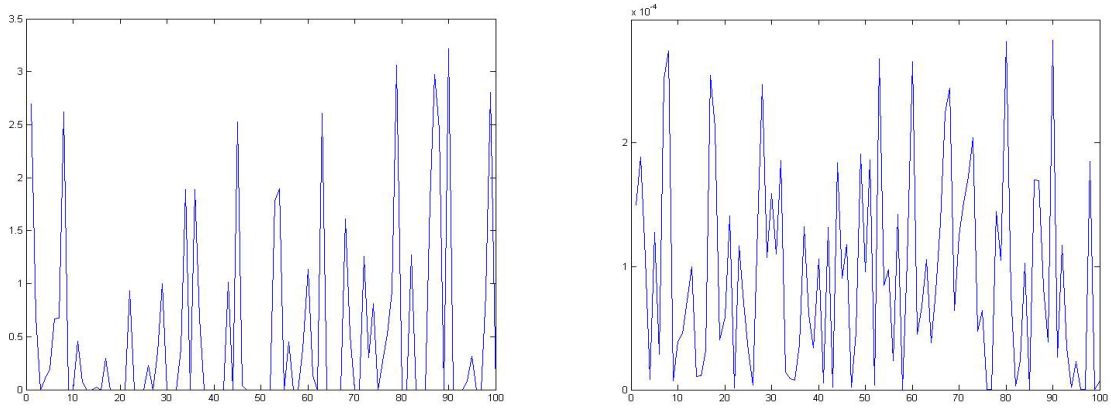


Figura 87 – Plot dos coeficientes H_{ij} para GP e MM

4.5.4.9 Qual o método NMF a seleccionar para este projecto?

Não foi fácil decidir qual o método mais favorável para o nosso trabalho, qual o número de

componentes e o número de iterações, em virtude de termos vários parâmetros que parecem ser difíceis de analisar, e com resultados de sinal oposto, tal como descrevemos e apresentámos na Tabela 3, na Tabela 4, na Tabela 5 e na Tabela 6.
A partir dos dados apresentados na secção anterior acabámos por optar para o nosso trabalho pelo método GP, com 100 componentes e 1000 iterações, nomeadamente para o estudo do sistema híbrido PCA+ICA+NMF.

4.5.5 NMF para imagens fora da BD

Em virtude de os eixos NMF não serem ortogonais, tal como acontecia com o ICA, uma imagem que não pertença à BD não pode ser decomposta nesse sistema por uma simples projecção ortogonal, calculando os coeficientes por produto interno dos vectores. Ou seja, somos uma vez mais obrigados a usar o algoritmo de decomposição *pesqproj* que descrevemos na secção 4.3.3.

Este processo tem uma contradição intrínseca, já que ao utilizarmos o algoritmo *pesqproj* não garantimos coeficientes positivos, o cerne do método NMF. No entanto parece-nos interessante avaliar se os eixos NMF, calculados segundo uma lógica distinta do PCA e ICA, conseguem melhorar o carácter esparso do processo de reconstrução, se associados ao PCA e ICA num sistema híbrido.

Tivemos de adaptar ligeiramente o algoritmo *pesqpro*, em virtude de termos no NMF eixos e vectores-imagem X com média diferente de zero, pois todos os valores das matrizes são por imposição positivos. Mas o algoritmo funciona como é descrito em 4.3.3, projectando iterativamente o vector erro nos eixos e escolhendo a projecção mais significativa. Começámos por tentar validar o algoritmo *pesqproj* da secção 4.3.3, modificado para o NMF, utilizando-o para decompor uma imagem interna à BD.

Utilizámos o sistema NMF seleccionado. Testámos o algoritmo *pesqproj* modificado para o NMF, com mil iterações, usando para projecção o sistema de eixos identificados pela matriz W, calculada pelo método GP, também com 1000 iterações e com 100 componentes. Obtivemos a curva de erro da Figura 88.

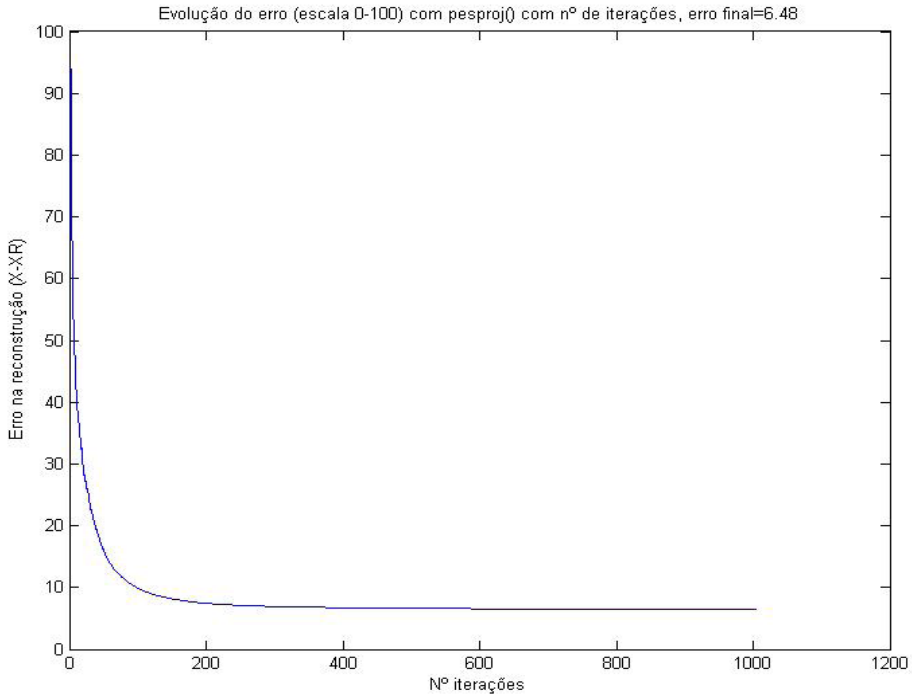


Figura 88 – Erro em função do nº iterações no algoritmo pesqproj – NMF(GP)

O erro final que calculámos é de cerca de 6.5 em 100, idêntico ao que obtivéramos directamente do NMF (ver Tabela 3), o que parece validar o algoritmo *pesqproj* para o NMF.

Utilizando os coeficientes H_i calculados pelo algoritmo *pesqproj*, em vez dos coeficientes calculados directamente pelo algoritmo NMF (GP), reconstruimos iterativamente o retrato,. Para isso utilizámos um script idêntico ao descrito no PCA e no ICA, para calcular iterativamente X_R com K componentes, K=1:100, recorrendo a um vector H_K, obtido de H fazendo zero os coeficiente de ordem> K.

$$X_R = W * H_K \qquad (55)$$

Obtivemos assim uma evolução para o erro representada na Figura 89. O valor final é idêntico (cerca de 6.5 em 100) o que valida mais uma vez o processo.

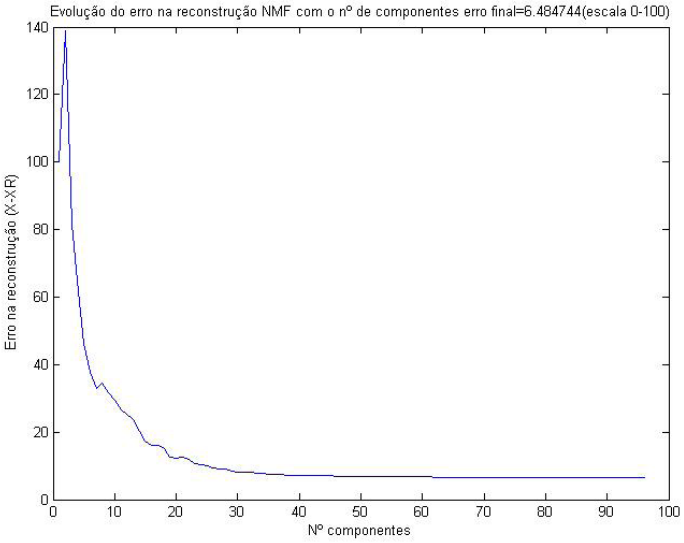


Figura 89 – Evolução erro com K, usando os coeficientes de projecção calculados por *pesqproj*

Testando uma imagem que não pertence à BD, obtivemos o erro representado nas Figura 90 e Figura 91, respectivamente com cada iteração do algoritmo *pesqproj* e reconstruindo iterativamente o retrato com K componentes NMF, utilizando os coeficientes calculados.

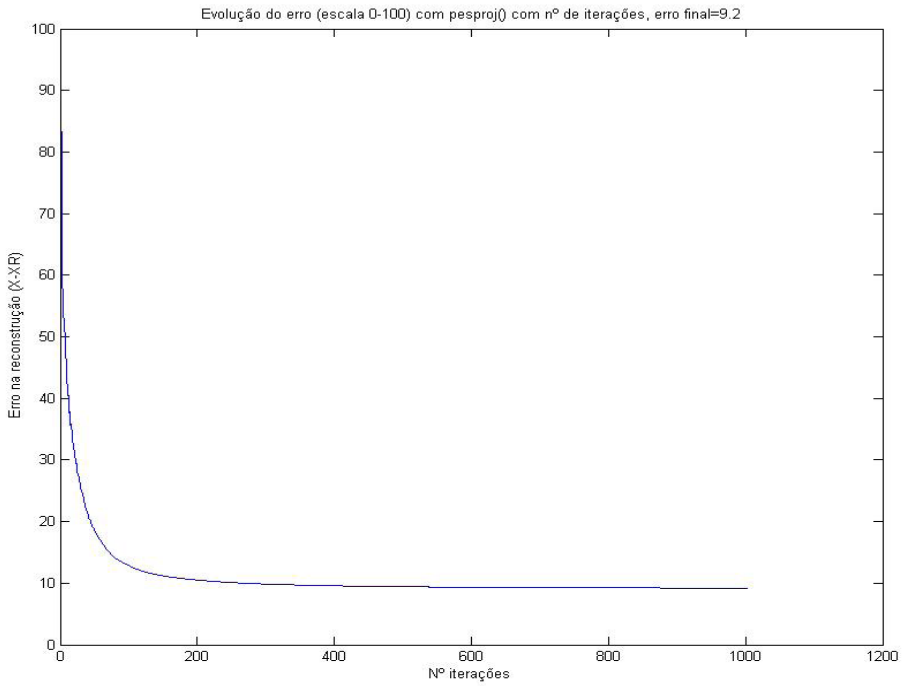


Figura 90 – Erro em função do nº iterações no algoritmo *pesqproj*

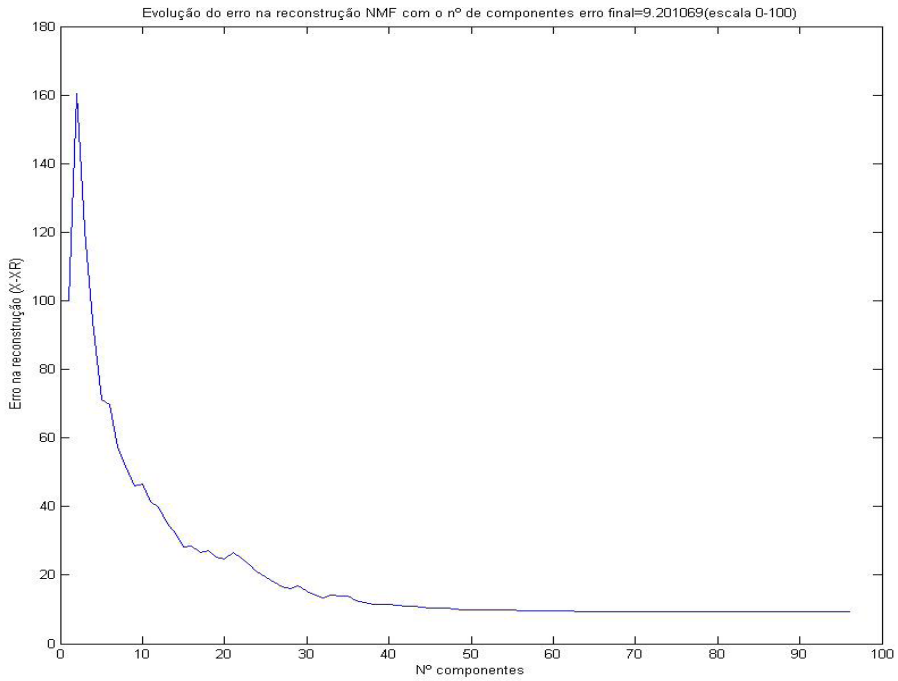


Figura 91 – Evolução erro com os coeficientes de projecção obtidos por pesqproj

Curiosamente em NMF o erro final obtido (9 em 100) para uma imagem que não pertence à BD não é muito pior do que o que se obtivera para uma imagem da BD, (cerca de 6.5 em 100). Ou seja, curiosamente, se o NMF não é comparável do ponto de vista da qualidade reconstrução de um retrato, ao PCA ou ICA, no caso das imagens da BD, onde esses métodos permitem uma reconstrução perfeita, o NMF permite obter um erro idêntico ao do PCA ou ao do ICA para imagens fora da BD, utilizando o algoritmo *pesqproj*.

Nota: não conseguimos explicar a subida inicial do erro com os primeiros componentes, depois de utilizar o algoritmo *pesqproj*.

4.6 Resultados com PCA + ICA + NMF

Vamos nesta secção tentar construir e testar um outro sistema híbrido para a nossa BD de retratos, criado com 900 eixos, associando os 400 componentes PCA, os 400 componentes ICA (*gauss*) e 100 componentes NMF (GP com 100 componentes).

Ao iniciar este trabalho fomos confrontados com uma dificuldade que parecia ser insuperável. De facto, quer o PCA, quer o ICA obrigam a preparar os vectores imagem, retirando a componente média. Os dois métodos trabalham com um vector X normalizado (XN), tal como referimos ao descrever o processo. O vector XN é o vector imagem X sem a componente média, como se mostra em (56)

$$XN=X-mean(X) \qquad (56)$$

Por isso, para decompor um vector imagem X num sistema PCA ou ICA, com o algoritmo *pesqproj*, temos de normalizar o X. Se não o fizermos, a reconstrução não se faz, pois o algoritmo não converge, ou seja as imagens parecem ficar fora do espaço N dimensional reconstruível pelo PCA/ICA.

Ora os algoritmos NMF obrigam precisamente ao contrário: a não poder retirar essa componente média, sob pena de termos em X valores negativos e pormos em causa a essência do método.

Conseguimos que o algoritmo *pesqproj* convergisse para o NMF decompondo o vector XN, a que o PCA/ICA obriga, para podermos construir o sistema híbrido que pretendemos testar, mas tal expediente resultou num erro superior. Em vez de um erro final de cerca de 6.5, que obtivemos com o *pesqproj* NMF, com a matriz X (que garante valores positivos, tal como o NMF impõe), obtivemos para o mesmo retrato, utilizando o algoritmo *pesqproj* com a matriz XN, um erro de cerca de 10 em 100. Fizemos várias experiências com médias e normalizações quer dos vectores W quer de X, para tentar obter um valor de erro idêntico ao de que partíramos, mas sem conseguirmos obter um valor da mesma ordem de grandeza.

A Figura 92 mostra a evolução do erro com o número de iterações do algoritmo *pesqproj* para o sistema NMF, utilizando a matriz XN. Podemos compará-la com a evolução do erro na Figura 88, que obtivemos com a matriz X.

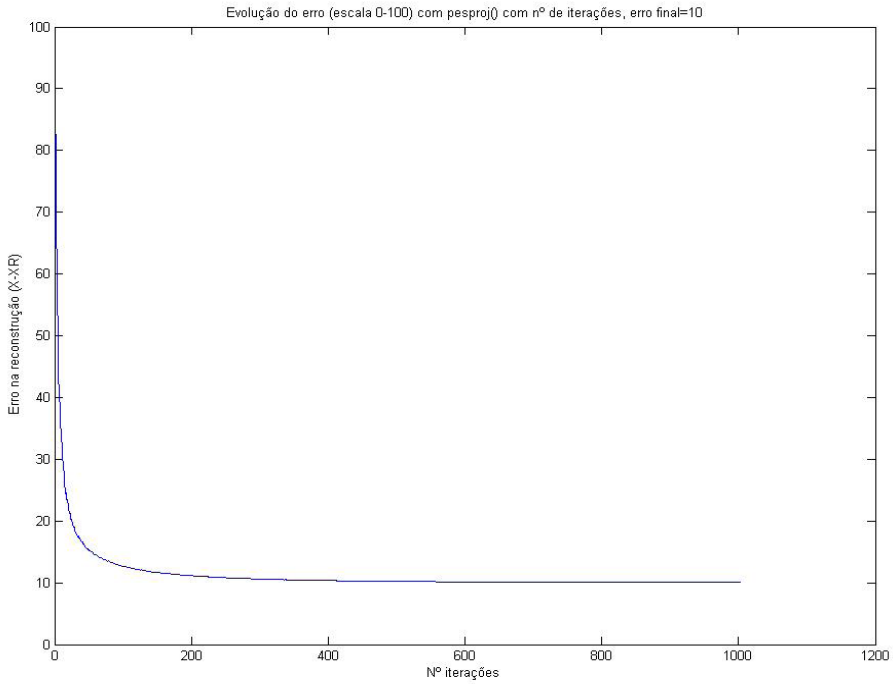


Figura 92 – Evolução do erro com as iterações em *pesqproj* utilizando XN

Esta diferença no erro sugere claramente uma questão de médias, que o algoritmo não consegue recuperar ao tentar reconstruir o vector, por utilizar a matriz XN em vez da matriz X.

Apesar de não conseguirmos recuperar o erro NMF inicial e de ser mesmo discutível a utilização do NMF, quando se põe em causa a sua essência, ao permitir no algoritmo *pesqproj* a obtenção de coeficientes não positivos, decidimos mesmo assim experimentar o sistema híbrido PCA+ICA+NMF.

Mas todos os testes que realizámos demonstram no entanto que o sistema POCA+ICA+NMF não parece acrescentar nada ao PCA+ICA. A evolução do erro com o número de componentes é semelhante, como se observa na Figura 93, e a 'esparsidade' não aumenta e por isso o número de componentes necessários é idêntico, como se vê na Tabela 7.

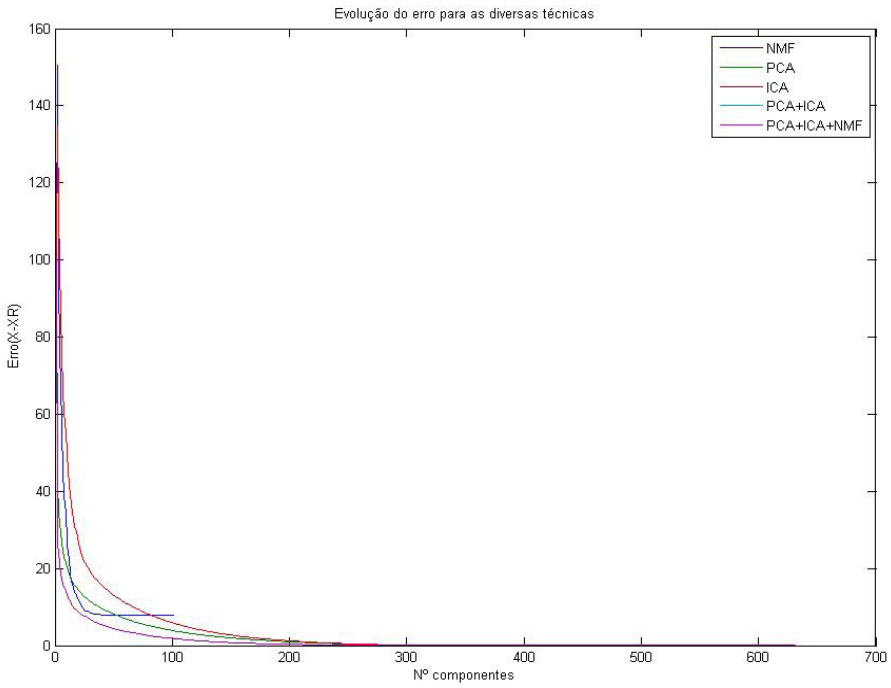


Figura 93– Evolução erro para as diferentes técnicas

A Tabela 7 compara o nº de componentes necessários para os dois sistemas híbridos, para garantir erros na reconstrução menores que 10, 1 e 0.1 numa escala de 0 a 100. Estes valores para o erro não são arbitrários como veremos, pois observa-se que o reconhecimento da imagem reconstruída se verifica com um erro entre 10 e 5 e que com erro < 0.1 a reconstrução é quase perfeita. Um erro > 1 corresponde a uma situação intermédia entre o limiar do reconhecimento e a reconstrução perfeita.

Método	**Erro < 10**	**Erro < 5**	**Erro < 1**	**Erro < 0.1**
PCA+ICA(*gauss*)	15	44	134	258
PCA+ICA+NMF	15	44	133	258

Tabela 7 – Limiares de erro para os sistemas híbridos para um retrato da BD

Verificamos portanto que se consegue reconstruir e reconhecer uma imagem da BD com pouco mais de 15 componentes, utilizando o método PCA+ICA ou PCA+ICA+NMF. O NMF parece não acrescentar nada ao sistema PCA+ICA.

4.7 Conclusões sobre as diferentes técnicas estatísticas

Depois de termos construído as ferramentas necessárias para utilizar todas as técnicas estatísticas com a nossa BD de retratos e depois de as termos validado e comparado, utilizando um retrato da BD e um retrato fora da BD, construímos os scripts Matlab necessários para avaliar o comportamento médio de cada um dos sistemas, incluindo os sistemas híbridos, tendo em conta todos os retratos da BD e o conjunto global dos retratos de teste. Estes scripts mais não fazem que aplicar os algoritmos descritos nas secções anteriores, retrato a retrato, a partir de uma directoria de retratos, e depois calcular os comportamentos médios, por exemplo a evolução média do erro com o número K de componentes utilizados na reconstrução.

Comparamos nesta secção os resultados obtidos com os vários métodos e comparamos também com os resultados da secção anterior, obtidos para um único retrato. Esta secção resume assim os resultados para as diferentes técnicas estatísticas avaliadas: o PCA, o ICA, o NMF e a solução híbrida PCA+ICA. Nesta comparação global excluímos o PCA+ICA+NMF por termos concluído, ao contrário das nossas expectativas iniciais, que o NMF não introduz benefício na reconstrução dos retratos.

A comparação é feita para os 400 retratos da BD e para os 39 retratos de teste que não pertencem à BD. Esta comparação vai considerar diferentes itens:

- o erro na reconstrução
- a relação entre o erro geométrico e a qualidade fotográfica
- o nº de componentes necessários para conseguir alguns erros
- a 'esparsidade'

4.7.1 Comparação para todos os retratos da BD

4.7.1.1 Erro na reconstrução

Reconstruímos os 400 retratos da BD de retratos um a um com cada uma das técnicas. Na Figura 94 pode-se observar o erro obtido para os 400 retratos da BD com o PCA: todas as curvas representadas a azul.

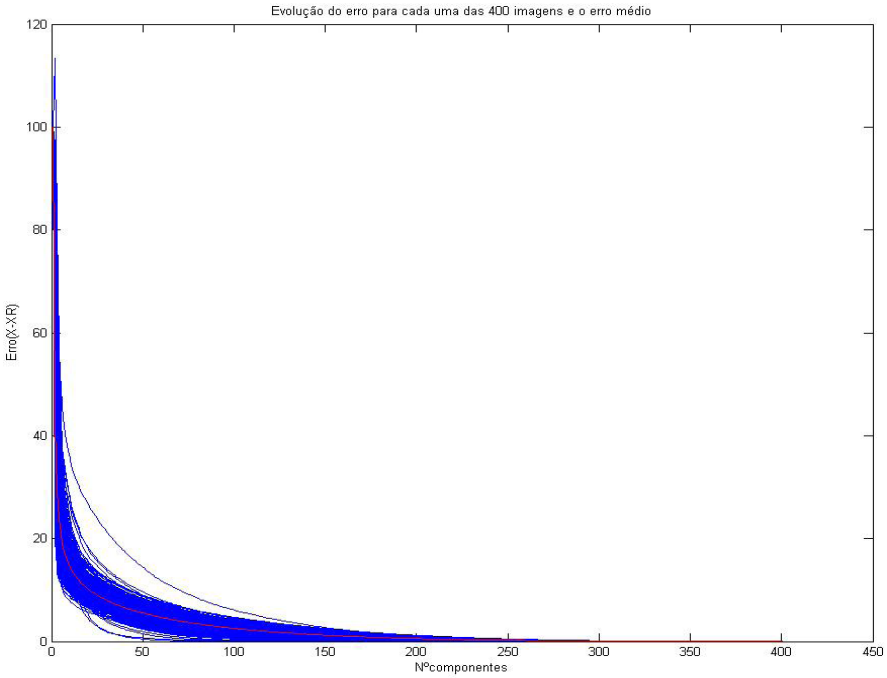


Figura 94 – Curvas erro de reconstrução obtidas para os 400 retratos com o PCA

A curva vermelha representa o erro médio associado à reconstrução com o PCA. As figuras seguintes mostram o mesmo resultado com o ICA, NMF e com o PCA+ICA.

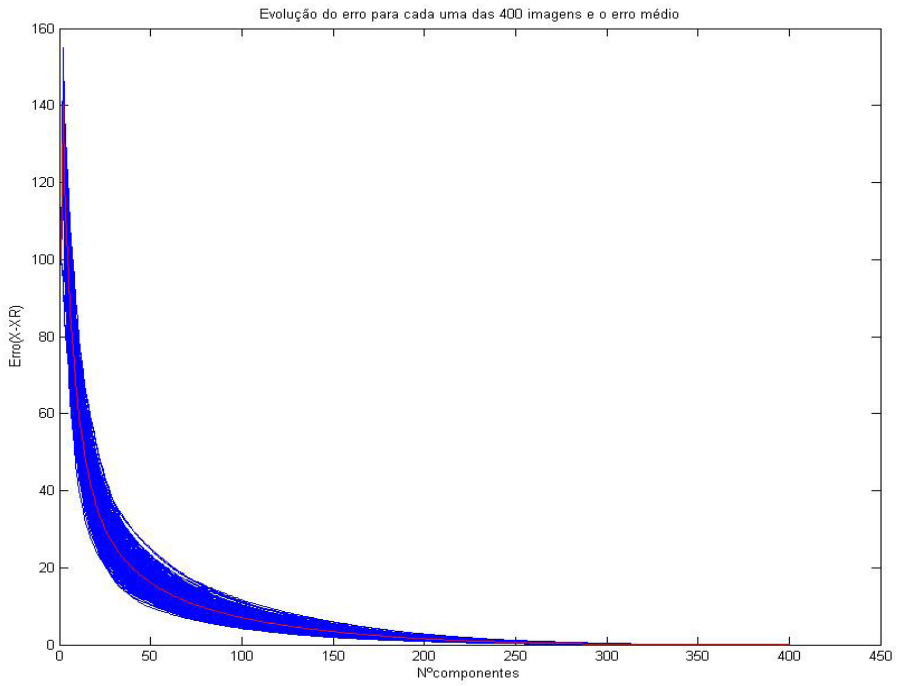


Figura 95 – Curvas erro de reconstrução obtidas para os 400 retratos com o ICA

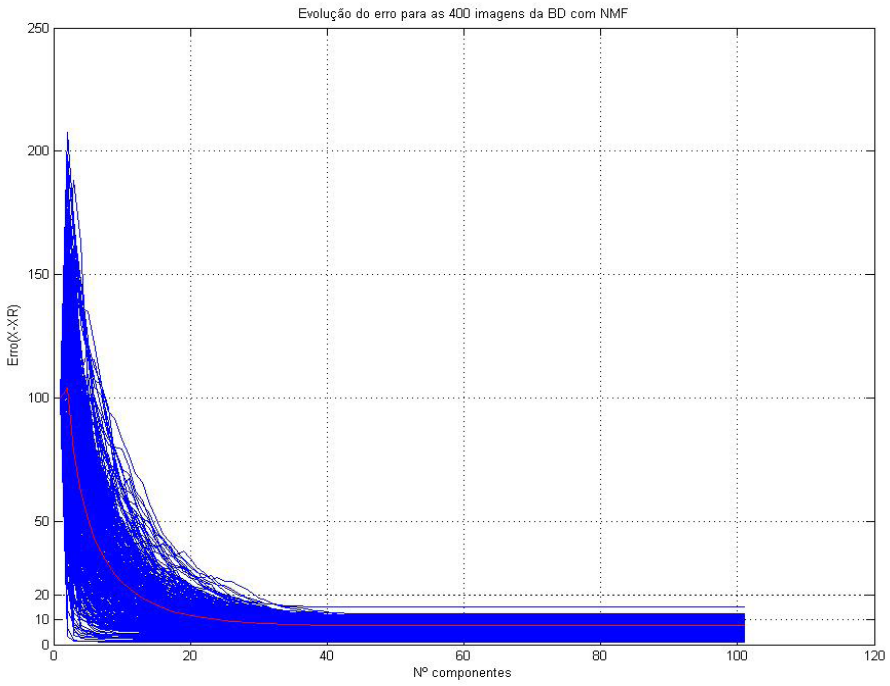


Figura 96 – Curvas erro de reconstrução obtidas para os 400 retratos com o NMF

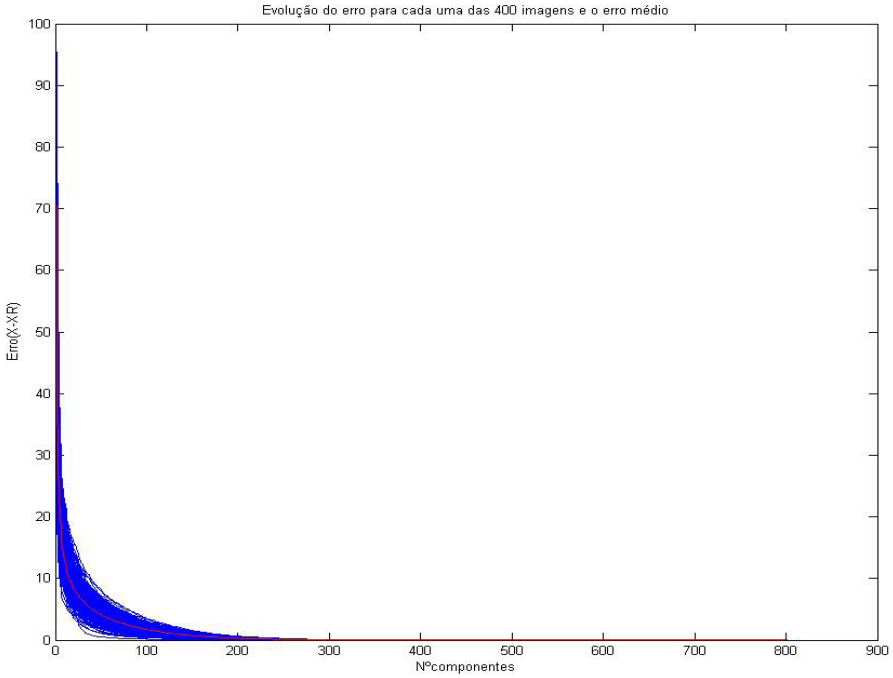


Figura 97 – Curvas erro de reconstrução obtidas para os 400 retratos com o PCA+ICA

Será interessante também comparar apenas a evolução do erro médio para os 3 casos em que o erro converge para zero: PCA, ICA e PCA+ICA.

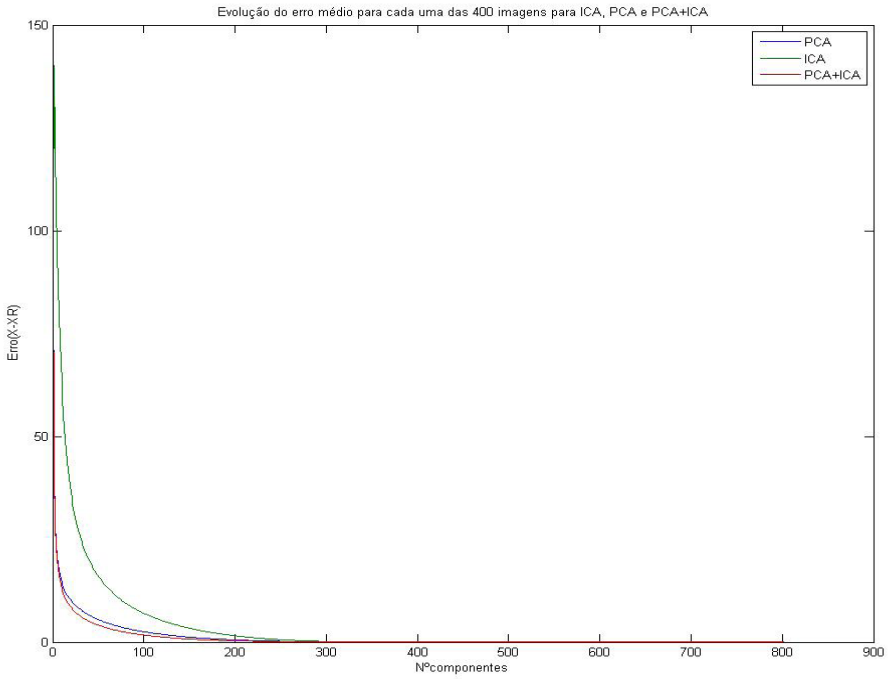


Figura 98 – Erro médio obtido para os 400 retratos com PCA, ICA e PCA+ICA

A Figura 98 mostra que a evolução dos erros médios para as 400 imagens para os 3 métodos não é muito diferente da que tínhamos observado com uma única imagem. Observamos erros que tendem todos claramente para zero. No PCA a curva converge um pouco mais rapidamente que no ICA, mas o sistema híbrido PCA+ICA é o mais favorável, pois converge ainda mais rapidamente. Este resultado é também visível na Tabela 8, observando o nº de componentes necessários à reconstrução do retrato com erros pré-fixados.

4.7.1.2 Número de componentes necessários

A partir dos erros médios apresentados da secção anterior, para cada uma das técnicas, calculados para o conjunto de 400 retratos da BD, a partir do erro calculado para cada retrato, é possível construir a Tabela 8, onde observamos o nº de componentes necessários, K, para garantir vários valores do erro médio, ao reconstruir o retrato com K componentes.

Técnica	Erro<12	Erro<10.5	Erro<9	Erro<7.5	Erro<6	Erro<4.5	Erro<3	Erro<1	Erro<0.1
PCA	K=13	K=18	K=23	K =33	K=44	K=61	K=88	K=161	K=281
ICA	K=65	K=73	K=81	K=93	K=110	K=129	K=156	K=223	K=315
NMF	K=18	K=22	K=27	Nota[3]	-	-	-	-	-
PCA+ICA	K=11	K=14	K=18	K=24	K=32	K=45	K=67	K=133	K=259

Tabela 8 – Limiares K de erro médio para retratos da BD para os vários métodos

Estes erros cobrem a zona interessante, desde valores onde o reconhecimento é difícil até onde o reconhecimento se faz com grande grau de certeza, como veremos na secção 6. Correspondem também aos valores utilizados nos inquéritos apresentados na secção 6. Veremos que com erros superiores a 10 em 100, o reconhecimento parece ser muito difícil, com erros inferiores a 0.1 o retrato recuperado parece ser quase perfeito e que entre 9 e 4.5 parecemos estar na zona crítica para um reconhecimento.

4.7.1.3 Avaliação das estatísticas dos coeficientes

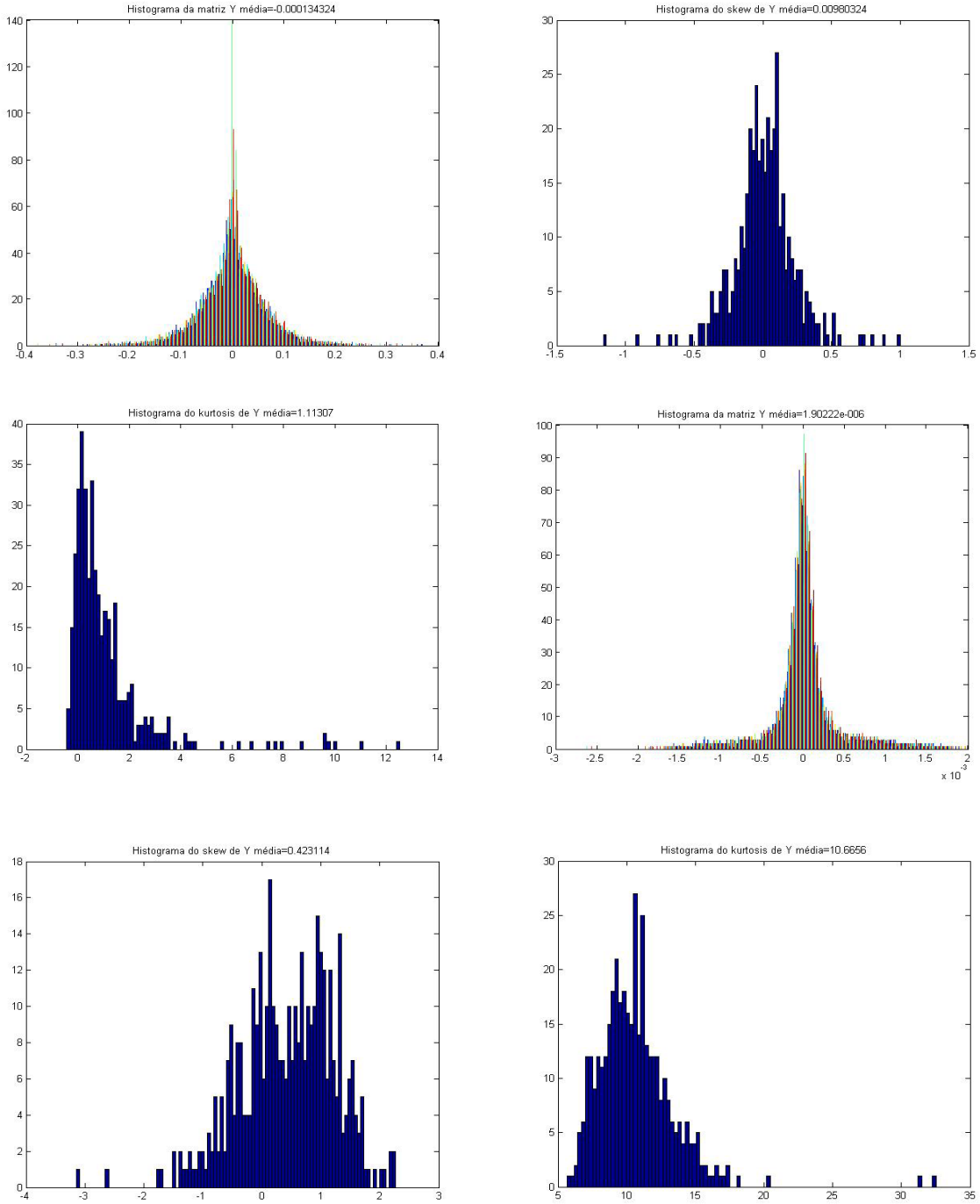
Para ajudar a analisar e a comparar os resultados obtidos para os coeficientes de mistura, calculados para cada uma das diferentes técnicas utilizadas para a nossa BD de retratos, mais uma vez analisámos a distribuição desses coeficientes e as estatísticas de 3ª e de 4ª ordem: a assimetria (*skewness*) e o 'achatamento' (*kurtosis*). Estas estatísticas podem ser associadas aos resultados obtidos anteriormente permitindo assim aprofundar a comparação entre as diferentes técnicas. Cada grupo de 3 histogramas na Figura 99 apresenta os resultados por nós obtidos, respectivamente para o PCA, ICA e PCA+ICA. Cada histograma foi calculado, considerando os valores das estatísticas, obtidas para os coeficientes de mistura, calculados para cada um dos 400 retratos da BD.

Os 3 histogramas para cada um dos métodos são:

- o histograma que representa a distribuição dos coeficientes em torno de zero
- o histograma da estatística de *skewness* para os 400 retratos
- o histograma da estatística de *kurtosis* para os 400 retratos

Será interessante fazer uma curta reflexão em torno dos resultados obtidos: como se sabe o paradigma de análise estatística é a distribuição normal, para a qual a assimetria é zero e o *kurtosis* 3. Todos os métodos calcularam coeficientes centrados em zero e concluímos que no PCA e no ICA a assimetria é também praticamente zero. Já no PCA+ICA existe uma assimetria nos coeficientes (-14) que poderá derivar da não ortodoxia formal deste método, que utiliza coeficientes PCA e ICA misturados, sendo portanto a simetria natural dos coeficientes afectada. Quanto à *kurtosis* observamos que o valor médio cresce de cerca de 1.11 para o PCA, para 10.67 no ICA e 343.45 no PCA+ICA. Estes valores podem talvez ajudar a compreender alguns resultados anteriores. No ICA o valor maior do *kurtosis*, que reflecte um maior 'achatamento' dos valores dos coeficientes em torno de zero, poderia levar-nos a esperar uma maior 'esparsidade' na reconstrução, ou seja uma reconstrução mais rápida do que no PCA. O facto de tal não acontecer dever-se-á ao reverso da medalha. O achatamento dos coeficientes em torno do zero é de facto maior, como

se observa, mas existem também coeficientes com maior valor, como se observa no histograma dos coeficientes do ICA. Este resultado permitirá no entanto talvez explicar a melhor qualidade fotográfica que o ICA parece oferecer para valores do erro elevados e que observámos em várias situações. O valor bastante elevado do *kurtosis* para o PCA+ICA parece estar associado neste caso de facto à maior 'esparsidade', que resulta da complementaridade PCA e ICA e que explica o resultado em que observámos uma reconstrução mais rápida dos retratos.



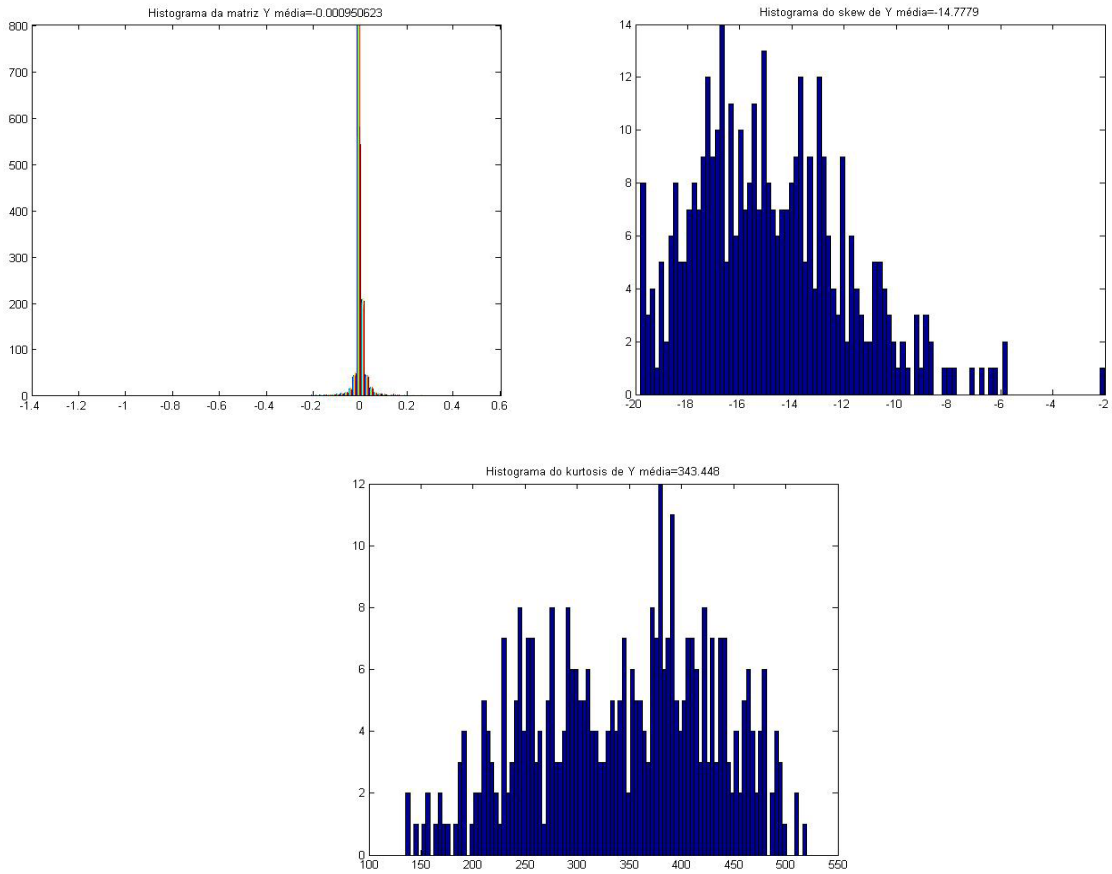


Figura 99 – Histogramas das estatísticas para o PCA, ICA e PCA+ICA

Temos consciência de que estas estatísticas, aqui como em situações anteriores, mereceriam porventura uma avaliação mais aprofundada do seu significado, que no entanto se não enquadra nos constrangimentos do horizonte temporal deste trabalho.

4.7.1.4 Erro geométrico e erro fotográfico

Para além dos valores do erro geométrico inerente ao processo iterativo de reconstrução de todos os métodos, que medimos em cada iteração do processo, utilizando a equação (48) e que apresentamos na secção anterior, interessa-nos também o erro fotográfico, que é um erro subjectivo que está relacionado com a qualidade da reconstrução fotográfica dos diferentes processos.

A Figura 100 representa a imagem reconstruída para o mesmo retrato, para valores do erro de 10, 5, 1, 0.1. As imagens de cada uma das linhas da figura são relativas a cada uma das técnicas testadas: PCA, ICA, PCA+ICA, PCA+ICA+NMF e NMF. No canto superior direito de cada imagem temos o erro de reconstrução associado. A coluna da direita mostra os originais. Observa-se que a qualidade fotográfica associada a um determinado valor do erro parece ser idêntica para todas as técnicas. Apenas no ICA, curiosamente, se observa que a qualidade fotográfica parece ser bastante melhor para um erro < 10. O caso do NMF é especial pois não se conseguem valores < 5.

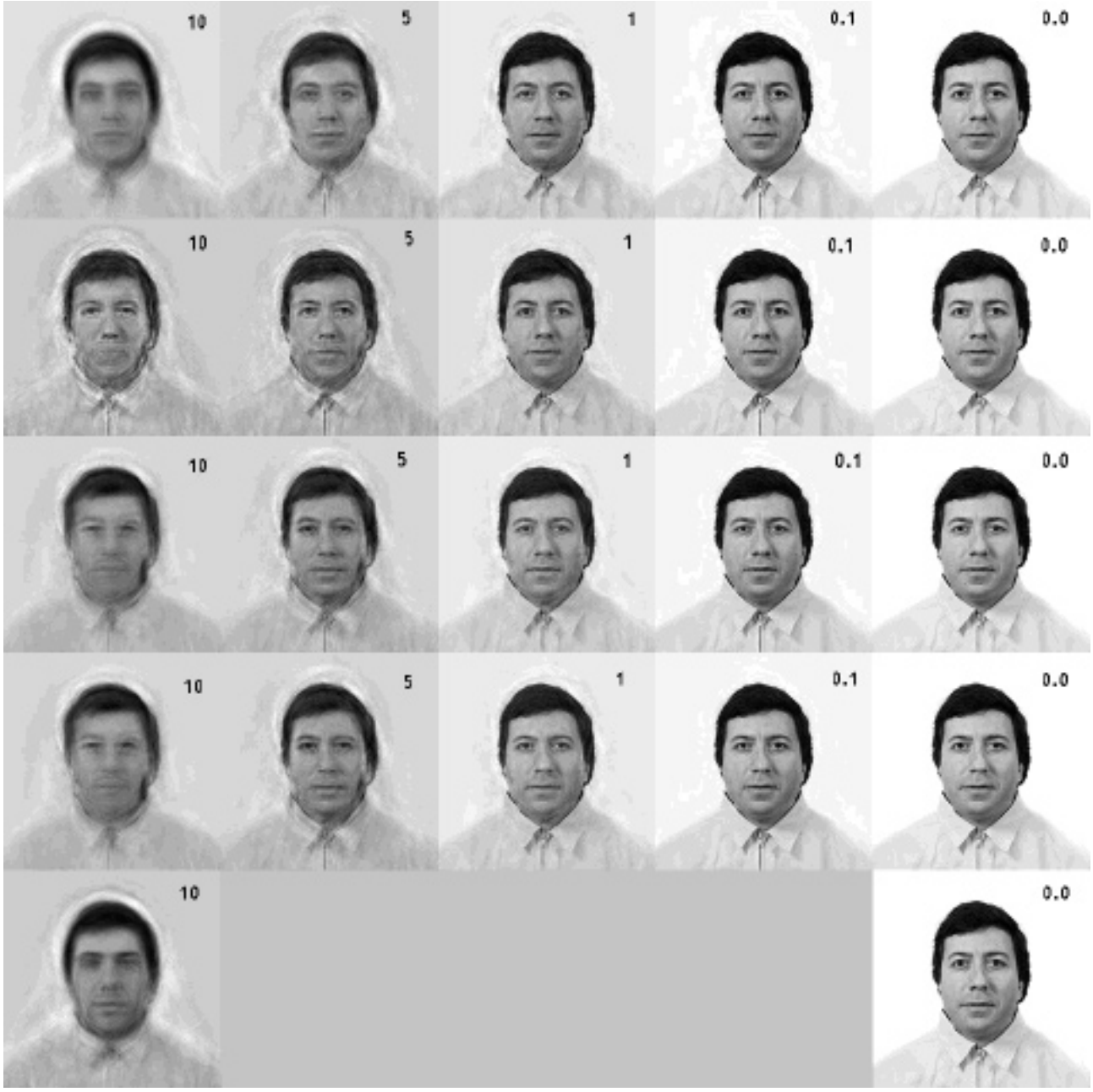


Figura 100 – Qualidade fotográfica em função do erro para um retrato da BD

Este resultado permite-nos constatar parecer haver uma correlação forte entre o erro geométrico e a qualidade fotográfica dos retratos reconstruídos. Ora a qualidade fotográfica parece ser essencial para um reconhecimento visual. Esta observação vai dar algum suporte a algumas das conclusões do inquérito que apresentamos na secção 6, onde pretendemos avaliar o reconhecimento fotográfico humano em função do valor do erro geométrico.

4.7.2 Comparação para todos os retratos fora da BD

Utilizando uma abordagem idêntica à que usámos para comparar os resultados para todos os retratos da BD, vamos aqui comparar os resultados obtidos para o conjunto de retratos fora da nossa BD. São 39 imagens que colocámos num grupo de testes, fora da nossa BD de 400 retratos. Mais uma vez não fizemos esta comparação global para o PCA+ICA+NMF, em virtude dos resultados obtidos anteriormente.

Mais uma vez tivemos de generalizar os scripts em Matlab que tínhamos escrito para fazer o cálculo retrato a retrato e calcular os valores médios, por exemplo para a evolução do erro de reconstrução com o número K de componentes utilizados. No caso do ICA e do PCA+ICA em virtude de serem sistemas não ortogonais tivemos de generalizar o algoritmo *pesqproj* para decompor o conjunto de todos os retratos de teste, calculando as matrizes A de coeficientes.

4.7.2.1 Erro na reconstrução

Reconstruímos os 39 retratos um a um com cada uma das técnicas. Na Figura 101 pode-se observar o erro obtido para os 39 retratos fora da BD com o PCA.

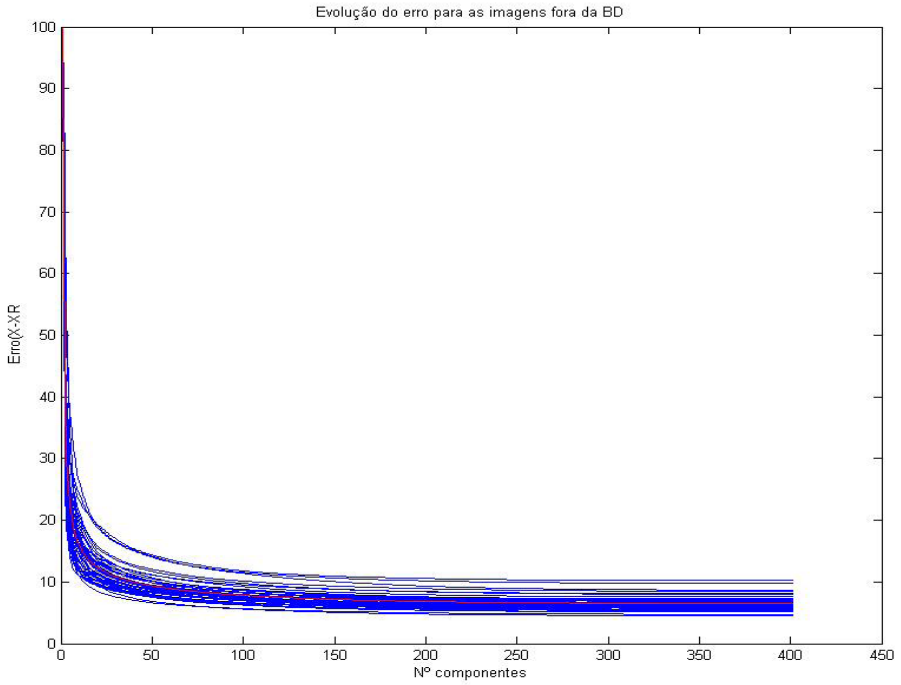


Figura 101 – Erros obtidos para os 39 retratos de teste com o PCA

A curva vermelha representa o erro médio associado à reconstrução dos retratos com o PCA. Obtém-se um erro final médio de 6.5 em 100 e um erro final mínimo e máximo de 4.5 e de 10.2 respectivamente.

As figuras seguintes mostram o mesmo resultado com o ICA, NMF e com o PCA+ICA.

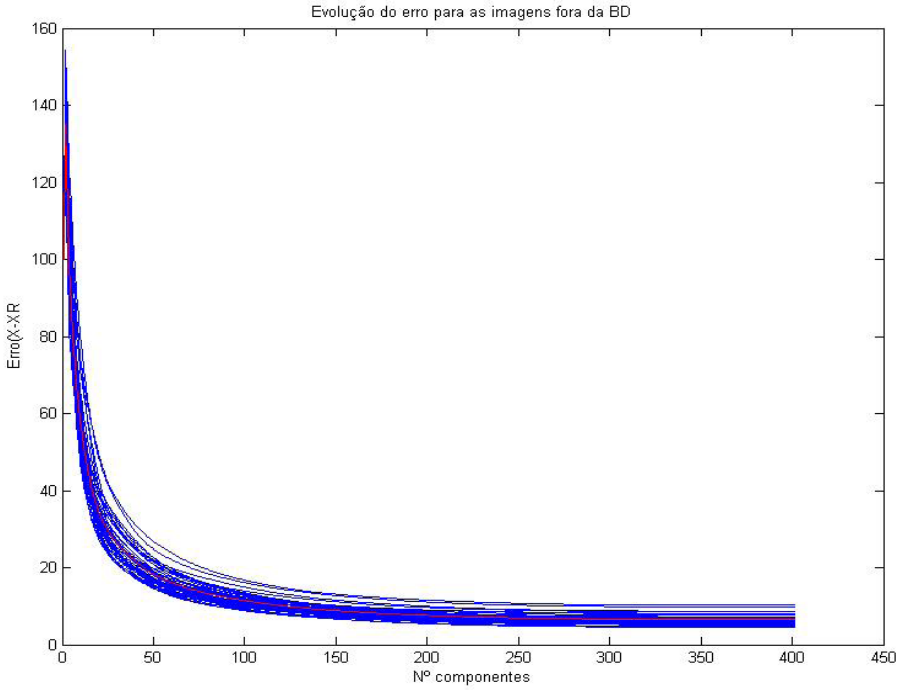


Figura 102 – Erros obtidos para os 39 retratos de teste com o ICA

Na Figura 102 a curva vermelha representa mais uma vez o erro médio, desta vez associado à reconstrução dos retratos fora da BD com o ICA. No caso do ICA observa-se que o erro final médio é igual ao do PCA, 6.5 em 100, e também os valores mínimos e máximos são os mesmos: 4.5 e 10.2.

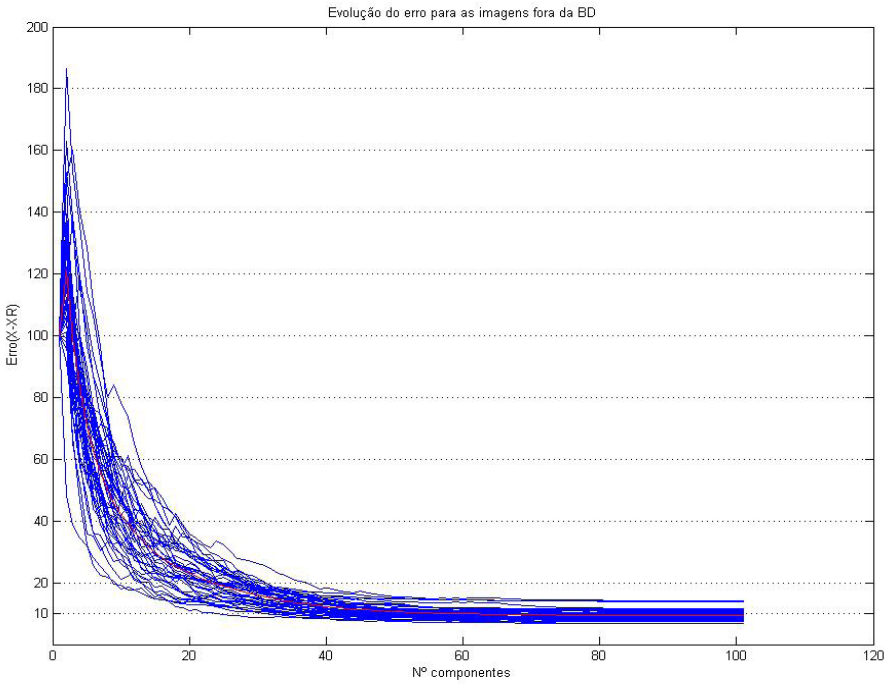


Figura 103 – Erros obtidos para os 39 retratos de teste com o NMF

Na Figura 103 a curva vermelha representa mais uma vez o erro médio, desta vez associado à reconstrução dos retratos fora da BD com o NMF (método GP com 100 componentes). No caso do NMF o valor final do erro médio é de 9.3 em 100 e os valores mínimo e máximo são neste caso: 6.8 e 14.2.

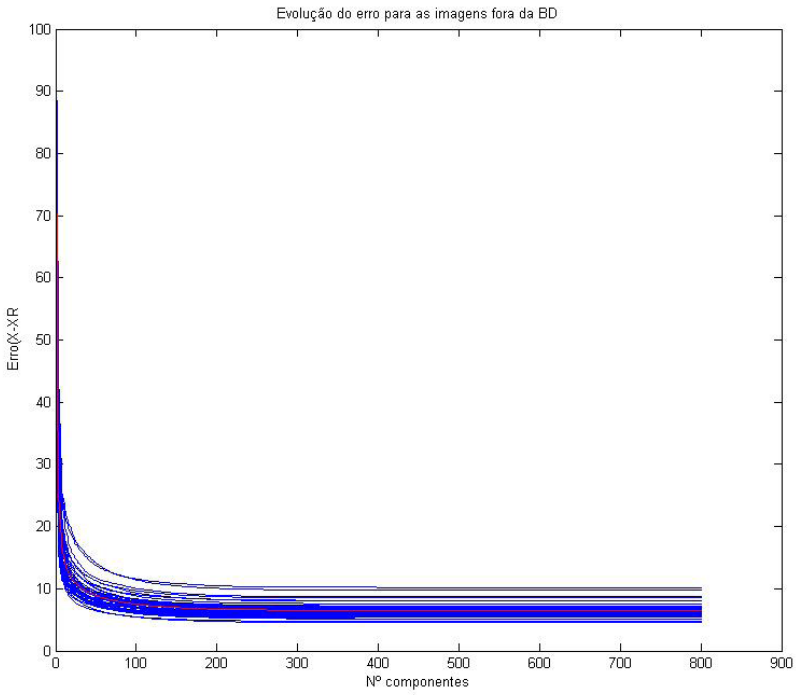


Figura 104 – Erros obtidos para os 39 retratos de teste com o PCA+ICA

Na Figura 104 a curva vermelha representa mais uma vez o erro médio, desta vez associado à reconstrução dos retratos fora da BD com o PCA+ICA. Observa-se que de facto apenas existem cerca de 600 componentes (dos 800 possíveis) com coeficientes diferentes de zero. Observa-se que o erro final médio é o mesmo, 6.5 em 100, e que também os valores mínimos e máximos do erro são os mesmos: 4.5 e 10.2.

Observa-se na Figura 105, que permite comparar os erros médios, e na Tabela 9 que a curva do erro médio para o PCA+ICA não difere muito da do PCA, mas a convergência é um pouco mais rápida no caso do PCA+ICA. Este resultado poderá traduzir de alguma forma o facto de o PCA ser porventura o método que melhor agrega os conhecimentos genéricos de retrato, adquiridos a partir de duma BD de N retratos, permitindo assim uma boa aproximação mesmo de retratos novos.

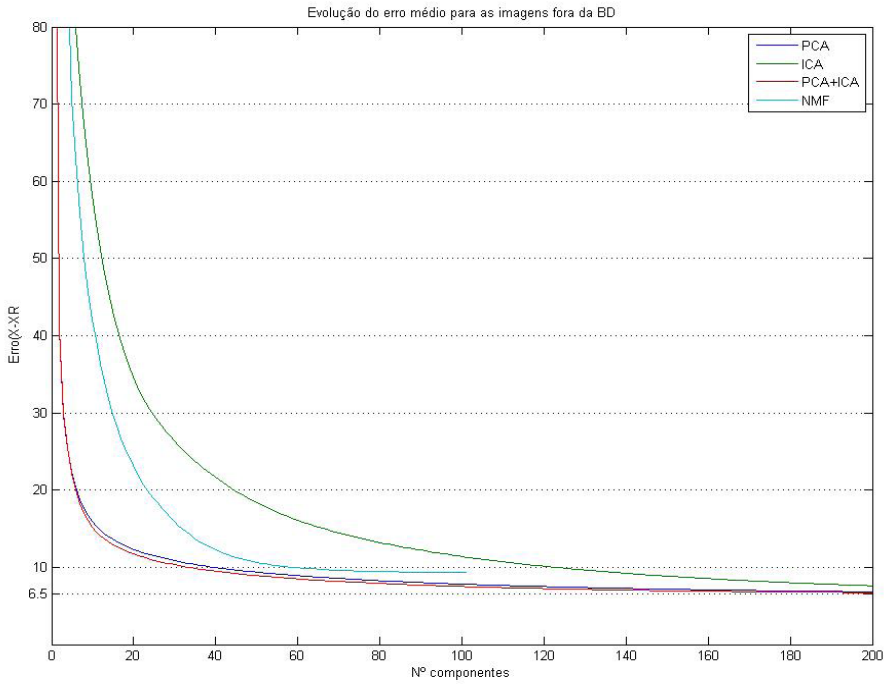


Figura 105 – Erros médios obtidos para os 39 retratos de testecom o PCA, ICA e PCA+ICA

4.7.2.2 Número de componentes necessários

A partir dos erros médios apresentados na secção anterior, obtidos para os 39 retratos fora da BD, para cada uma das técnicas, é possível construir uma tabela onde observamos o nº de componentes K necessários para garantir vários valores do erro.

Técnica	Erro Final	Erro<12	Erro<10.5	Erro<9	Erro<7.5	Erro<6
PCA	6.5	K=21	K=33	K=57	K=120	-
ICA	6.5	K=92	K=113	K=145	K=205	-
NMF	9.3	K=41	K=51	-	-	-
PCA+ICA	6.5	K=18	K=28	K=48	K=100	-

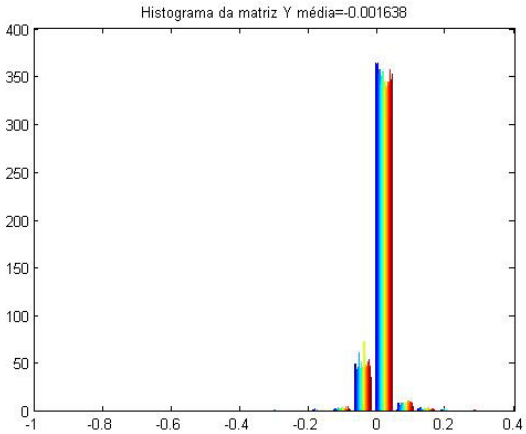
Tabela 9 – Limiares de erro médio para retratos fora da BD para os vários métodos

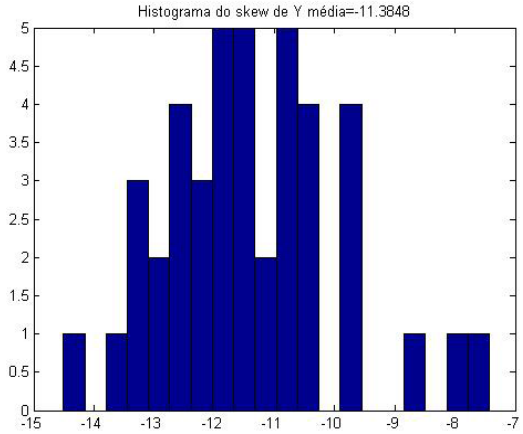
Observamos mais uma vez que o PCA, apesar de ter erros finais idênticos, permite uma reconstrução muito mais rápida do que o ICA. O PCA+ICA permite uma reconstrução ainda um pouco mais rápida.

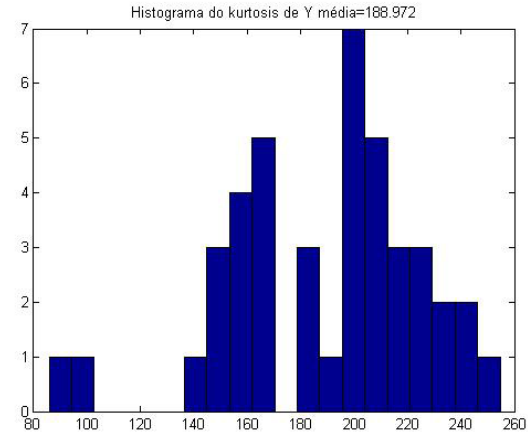
Os erros escolhidos cobrem um subconjunto dos valores de erro também utilizados na tabela idêntica (Tabela 8), relativa à reconstrução dos retratos da BD, e é um subconjunto por não se conseguirem valores de erro inferiores ao erro final da reconstrução.

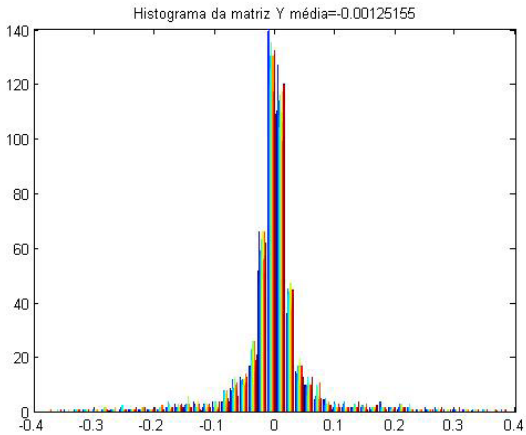
4.7.2.3 Avaliação das estatísticas dos coeficientes

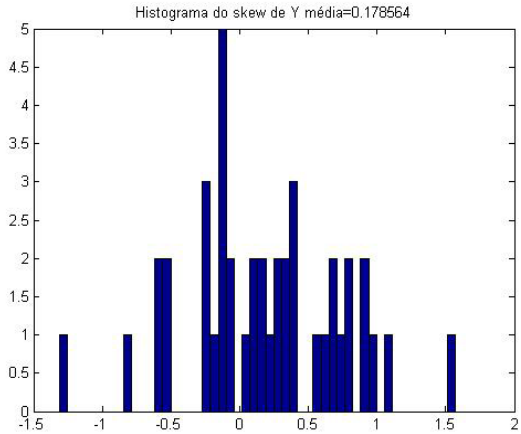
Para analisar e comparar os resultados calculados para os coeficientes de mistura, para as diferentes técnicas utilizadas, neste caso para as imagens fora da nossa BD, analisámos mais uma vez a dispersão desses coeficientes em torno de zero e comparámos as estatísticas de assimetria (*skewness*) e de 'achatamento' (*kurtosis*). Os coeficientes voltam a estar centrados em zero mas verificámos que a assimetria se afasta de zero no PCA. Os histogramas seguintes apresentam os 3 histogramas por nós obtidos, respectivamente para o PCA, ICA e PCA+ICA, representando para as estatísticas de assimetria e de 'achatamento' os valores médios calculados para os coeficientes dos 39 retratos de teste.

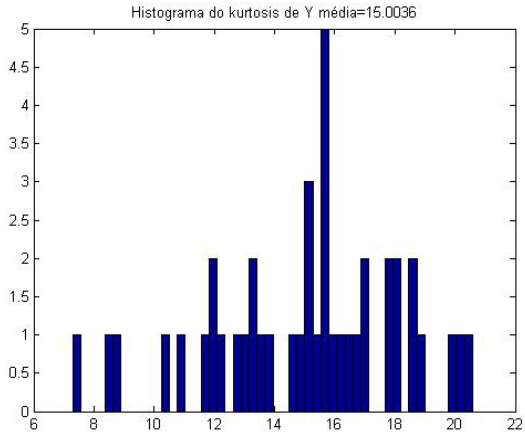












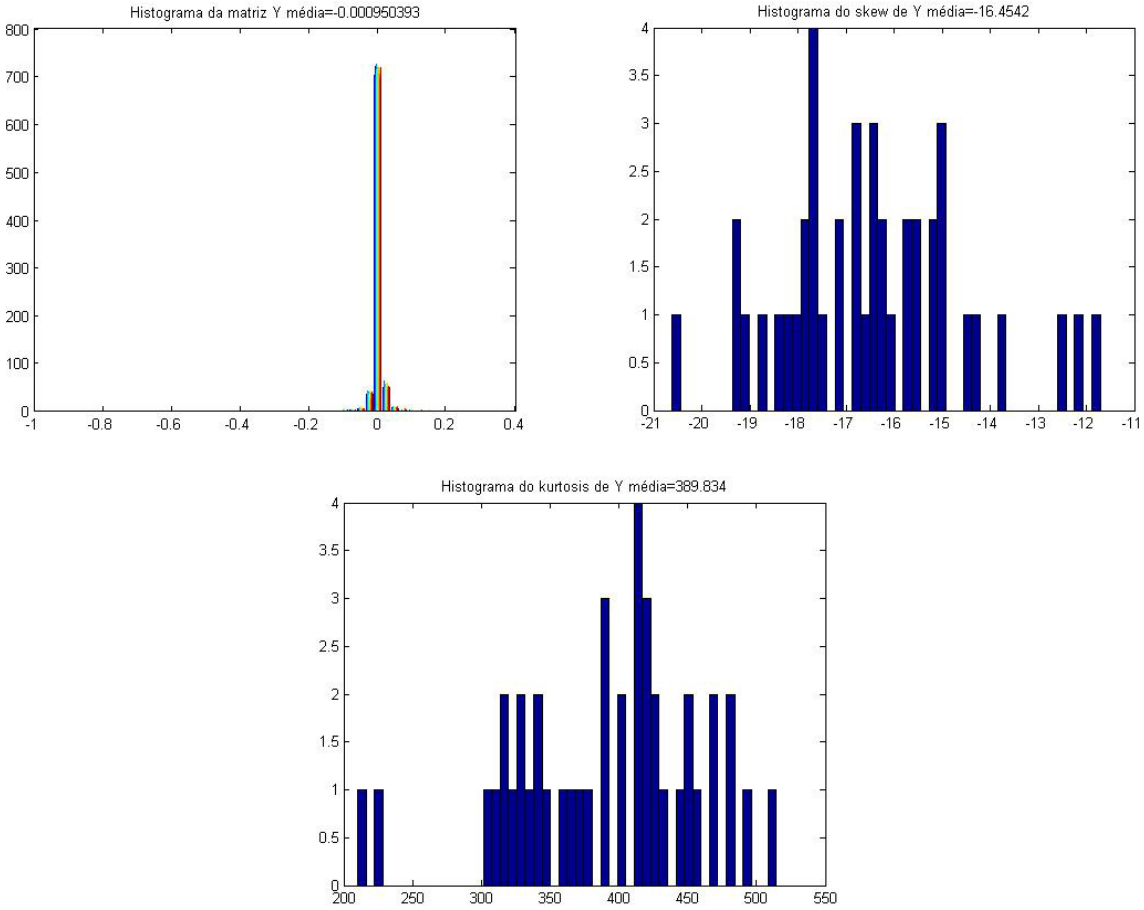


Figura 106 – Histogramas das estatísticas para o PCA, ICA e PCA+ICA

4.7.2.4 Erro geométrico e erro fotográfico

Como vimos o erro fotográfico é um erro subjectivo, relacionado com a qualidade da reconstrução fotográfica dos diferentes processos. A Figura 107 representa a imagem reconstruída para o mesmo retrato, que não pertence à BD, para um erro de 10 e para o erro final conseguido. As imagens em cada uma das linhas da figura referem-se a cada uma das técnicas testadas: PCA, ICA, NMF e PCA+ICA. Observa-se que a qualidade fotográfica associada a determinados valores do erro parece também neste caso ser idêntica para todas as técnicas, excepto mais uma vez no ICA onde o erro de 10% parece corresponder a uma melhor imagem.

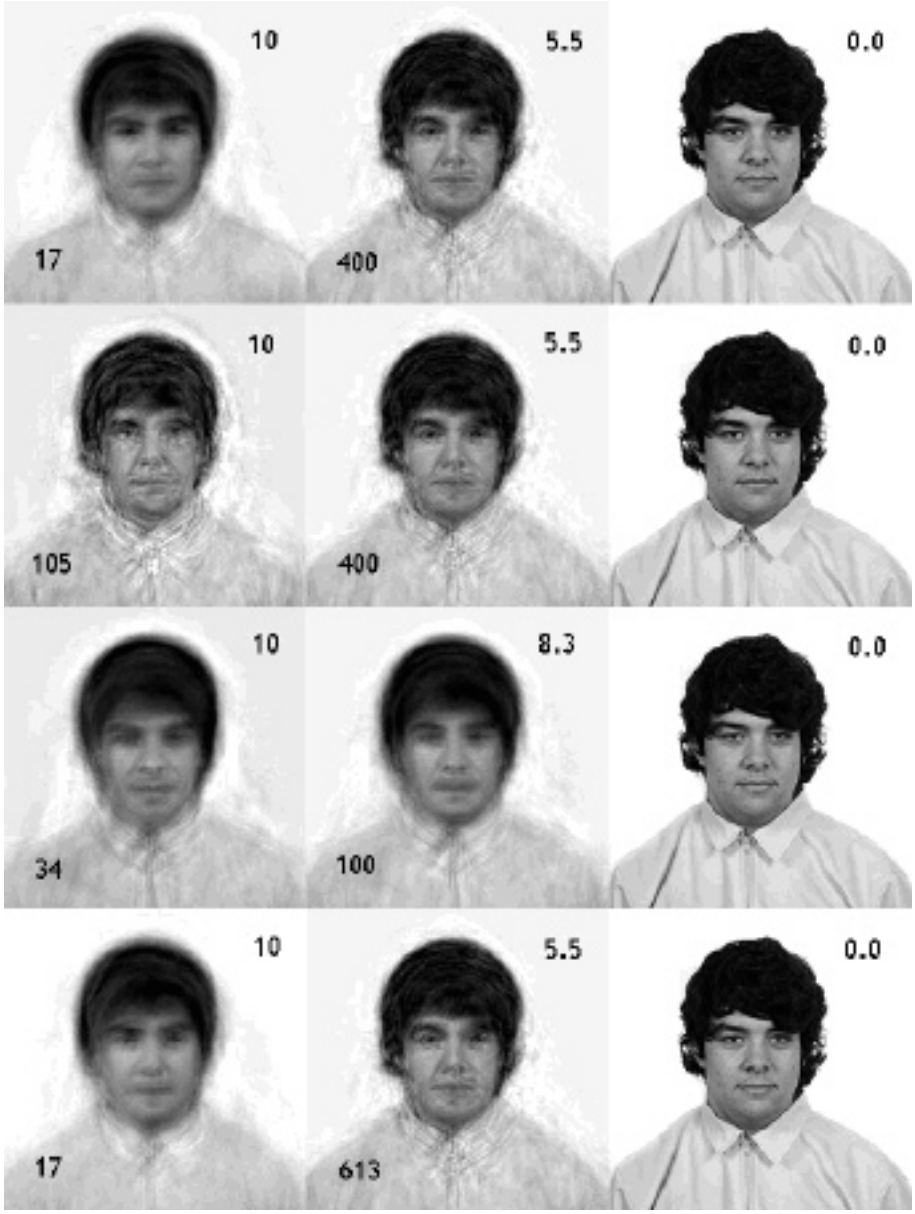


Figura 107 – Qualidade fotográfica em função do erro para um retrato fora da BD

No canto superior direito de cada retrato temos o erro na escala de 0 a 100 e no canto inferior esquerdo o nº de componentes necessários. A coluna do meio apresenta a reconstrução com todos os componentes, mas como se observa na Figura 105 para o PCA, ICA e PCA+ICA o erro pouco evolui a partir de 200 componentes. Os retratos da coluna da direita são os originais.

De realçar a qualidade, apesar de tudo, da reconstrução de uma imagem nova, que parece continuar a permitir um reconhecimento sem grandes dificuldades.

4.7.3 Erro fotográfico em função da resolução

Como vimos em secções anteriores para além do conceito de erro geométrico no processo de reconstrução de uma imagem, interessa-nos também o conceito de erro fotográfico, que ao contrário do erro geométrico é um conceito subjectivo e não mensurável e que está relacionado com a qualidade da reconstrução fotográfica dos diferentes processos. Interessa-nos validar se existe uma relação entre os dois valores, ainda que o segundo, como vimos, seja subjectivo.

Interessa-nos para isso também avaliar a possível modificação dessa relação entre o valor do erro geométrico (o único mensurável) e a qualidade fotográfica, se variarmos a resolução das imagens. Será legítimo, por exemplo, afirmar sempre que um erro de 5 em 100 possibilita um reconhecimento razoável, ou que erro de 1 em 100 parece não trazer qualquer dificuldade no reconhecimento, para qualquer resolução das imagens? Poderemos extrapolar sem problemas os resultados obtidos no inquérito da secção 6, com retratos de 200x200 pixéis, quanto à facilidade de reconhecimento, para outras resoluções?

Utilizando os nossos scripts de Matlab experimentámos reconstruir o mesmo retrato com o PCA, utilizando BDs com resoluções diferentes. Testámos a mesma BD de 400 retratos, que usámos ao longo do nosso trabalho, mas com uma resolução de 100x100, e uma nova BD com apenas 100 retratos[15] com uma resolução de 400x400. A Tabela 10 apresenta o número de componentes necessários para reconstruir um retrato com os limiares de erro pretendidos, em cada um dos casos.

Método	**Erro < 10**	**Erro < 5**	**Erro < 1**	**Erro < 0.1**
PCA (200x200) – 400 componentes	23	69	181	298
PCA (100x100) – 400 componentes	17	53	173	289
PCA(400x400) – 100 componentes	12	24	55	83

Tabela 10 – Limiares de erro no PCA para várias resoluções nos retratos da BD

Com base nos resultados da Tabela 10 podemos comparar a qualidade fotográfica das imagens associadas a esses valores do erro, para cada uma das resoluções: 200x200 / 100x100 e 400x400.

A primeira linha na Figura 108 apresenta os retratos reconstruídos com os erros indicados no canto superior direito para uma resolução de 200x200, a segunda linha apresenta os retratos reconstruídos com o mesmo erro, mas com uma resolução de 100x100 pixéis e a 3ª linha apresenta os retratos reconstruídos com o mesmo erro, mas com uma resolução de 400x400 pixéis.

15 O número de imagens e de componentes igual a 100 com 400x400 pixéis foi-nos imposto por limitações de memória no cálculo do PCA. Valores da resolução inferiores a100x100 tornam difícil avaliar a qualidade fotográfica.

Figura 108 – Qualidade fotográfica em função do erro para um retrato da BD

Apesar das imagens serem diferentes parece-nos que as diferenças de qualidade para um mesmo valor do erro geométrico não são perceptíveis. Parece portanto ser legítimo associar os valores do erro geométrico às qualidades relativas da reconstrução fotográfica, independentemente da resolução.

4.7.4 Erro fotográfico em função da dimensão da BD

Procurámos também analisar a evolução da qualidade fotográfica do retrato reconstruído, numa avaliação subjectiva, comparando retratos reconstruídos com BDs de dimensões diferentes, e comparando o número de componentes necessários[16]. A Figura 109 mostra o retrato reconstruído com um valor para o erro idêntico, com BDs de dimensão 50, 100, 200 e 400.

Figura 109 – Reconstrução com 18, 31, 52, 80 componentes, com BD de dimensão 50, 100, 200 e 400

A qualidade fotográfica parece uma vez mais idêntica, confirmando também neste caso a relação entre erro geométrico e qualidade fotográfica.

16 De notar que na secção anterior, no caso da BD com retratos 400x400, já tínhamos também uma alteração da dimensão da BD, imposta pelas restrições de memória.

Tal como no PCA procurámos também no ICA analisar a evolução da qualidade do retrato reconstruído, com BDs de dimensões diferentes, numa avaliação subjectiva, comparando os retratos reconstruídos com o mesmo erro geométrico para várias dimensões da BD, e comparando o número de componentes necessários.

Figura 110 – Reconstrução com 20, 40, 65 e 100 componentes, com BD de dimensão 50, 100, 200 e 400

A qualidade fotográfica parece, tal como no PCA, ser idêntica, confirmando também para o ICA a relação entre erro geométrico e a qualidade fotográfica, independentemente da dimensão da BD de retratos.

4.7.5 Limiares da dimensão da BD em função da resolução

Será ainda interessante completar a análise que fizemos na secção 4.2.2 e observar a dimensão da BD que é necessária, para cada resolução dos retratos, para garantir a qualidade fotográfica na reconstrução de retratos novos, para conseguir fazer um reconhecimento. De acordo com os resultados da secção 6, o reconhecimento crítico situa-se entre erros de 9 e 4.5, conseguindo-se com 9 um reconhecimento efectivo de cerca de 50% e com 4.5 um reconhecimento quase perfeito. Com base na Figura 50 podem-se obter esses limiares: N1 e N2. A Tabela 11 mostra os valores que encontrámos.

Resolução	**N1 (erro=9)**	**N2 (erro=4.5)**
200x200	177	600 (valor estimado)
100x100	125	350
50x50	60	180
20x20	-	40

Tabela 11 – Variação dos limiares da dimensão de uma BD com a resolução

Poderíamos estimar um número de cerca de 600 retratos para garantir um erro de 4.5 com a nossa BD 200x200 pixéis. A Figura 111 mostra a evolução da curva do limiar N2 com a resolução da BD, ou seja podemos estimar a dimensão da BD para assegurar um erro < 4.5 e consequentemente um reconhecimento quase perfeito. A curva tem um crescimento linear achatado.

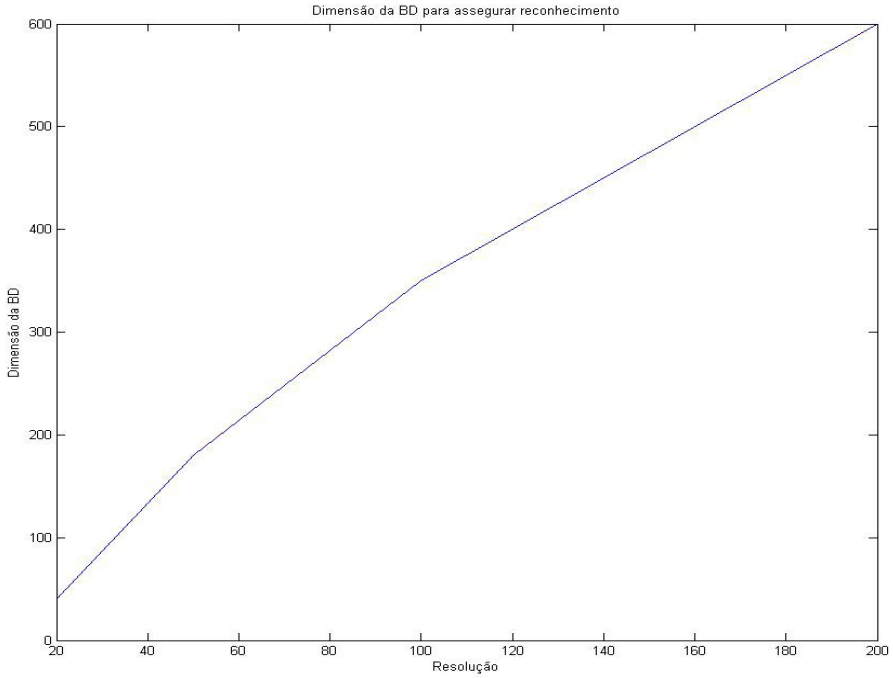


Figura 111 – Dimensão da BD com retratos de NxN pixéis para assegurar erro < 4.5 em retratos novos

Mas não acredita na minha identidade? Aqui tem a minha fotografia!

Caricatura alemã, 1877

Mnemodisseia

Todas as noites dobrava a roupa sobre a cadeira. Depois retirava com cuidado as sobrancelhas, os olhos, a boca, os sinais… até dispor o rosto em cima da escrivaninha, vigiado pelo espelho de pescoço esticado e moldura de pau-santo. Apreciava o sono despojado mas sentia haver outra razão para se deitar subtraído a si mesmo. Talvez o medo de ser reconhecido nessa imprevisível região dos sonhos, sabe-se lá por quem e com que implicações. Preferia evitar na memória todo e qualquer traço desse voo ou cavalgada nocturna.

Suspensa a *indumentária*, tacteava o caminho para a cama. De manhã fazia o mesmo, ao contrário. Recomeçava ordenadamente, como quem segue o manual instruções. Mas, uma vez colocado o *olho director*, a dúvida sondava. E agora, a boca ou o nariz? E a seguir?... Desencadeava-se o jogo. Parodiava os ciclopes (*e se saísse à rua de olho na testa?*), o merceeiro do quarteirão (*e esta sobrancelha no lugar do bigode a pente fino?*). Por fim normalizava o semblante, parodiando-se a si próprio. E saía, afinando o passo, na rua, pela certeza das aparências.

Para se especializar no seu despojo diário recorreu primeiro aos melhores *tupperware* encontrados na cozinha. Depois folheou catálogos, correu o comércio. Resignou-se, por fim, a gastar um pouco mais na encomenda. Considerou engenheiros, ourives, cientistas, mas decidiu-se pelo agente funerário (quem trata assim dos mortos merece a confiança dada em excesso a quem trata dos vivos). Fez a ginástica necessária para esconder a função e foi bem sucedido. Chegou a empalidecer quando o agente falou a despropósito nos romanos, nas máscaras, etc. (erudição comercial, concluiu). Levou para casa uma espécie de guarda-jóias capaz de arquivar o seu rosto até à mais pequena erupção cutânea.

Um dia acordou, pela primeira vez, com a lembrança do sonho que tivera: barulho de portas e gavetas. Mas, quando sentiu que uma brisa andava à solta como se a casa estivesse escancarada, correu à escrivaninha e apalpou um vazio. Agitou desordenadamente os braços porque não podia arregalar os olhos, deu murros sem poder gritar nem franzir o sobrolho. Correu a casa como se percorresse a palma da mão. E o pior foi depois da porta, escada abaixo e rua fora: porque não via nada, porque os outros o viam sem lhe ver as feições. À sua passagem ouviu gritos, exclamações, comentários. Foi amparado, empurrado, conduzido. Caiu, bateu, rodopiou, desceu. Chegou, por fim, à frieza e silêncio de um sítio, digamos, final. De pernas quebradas e peito arquejante, escorregou pela parede fria até à rigidez do chão. Ouvia rumores do outro lado da parede. “Entra”, disse-lhe uma voz congelada. Sentaram-no e deram-lhe cara nova, esculpindo e pintando, sem dor, uma espécie de prémio que não entendia. “Não te voltes”, disseram-lhe no fim, abandonando a sala. Pelo espelho viu dois desconhecidos. O primeiro, ele próprio, pois não se reconheceu. O outro, pouco diferente de si mesmo (daquilo que era agora): jovem, de traços finos, como que desenhados por uma eternidade bem afiada. Sentava-se numa escrivaninha igual à sua e remexia uns papéis sem deixar marcas. Mais atrás, numa parede impecavelmente pintada, uma tela perturbadora. Tudo nela era tinta, textura, empaste, raspagem, escorrido. Aos poucos foi percebendo uma forma humana descarnada. “Um retrato meu…”, disse o outro, adivinhando os pensamentos dele “… que mais valia estar, como os outros, enterrado aí em qualquer museu”. “Mas, então, tu és…”, disse, com o dedo indicador a procurar no vazio. “Nem digas!”, ordenou o outro. “Então é para isso que aqui estou?”, perguntou, “para um retrato? Uma pintura?”, inquiriu sem tirar os olhos do espelho. “Não, uma fotografia”, disse-lhe o outro (como se já ninguém quisesse pintar nem ser pintado).

A luz ofuscante do *flash* devolveu-lhe a rua. Começou a andar a passo lento como quem precisa de tempo para fazer bem as perguntas. Acabou por regressar, caminhando como se a casa estivesse do outro lado do mundo. Sentado à escrivaninha, tocou a face com a ponta dos dedos, experimentou retirá-la. Mas era impossível removê-la. Levantou-se e começou a andar em círculos, parando de vez em quando para olhar o espelho. Quis desenhar no semblante o ar grave que lhe enchia os pulmões mas nada se alterava, nem uma prega, nem um sinal de afectação. Tocou o telefone. Tinham encontrado o suspeito, era chamado a identificá-lo. Ocorreu-lhe dizer que não o tinha visto, que dormia durante o assalto, etc., mas achou melhor calar-se. Entrou no departamento e deu de caras consigo próprio, quer dizer, com alguém com as suas feições, a ser fotografado de frente, de lado, de costas. E, quando teve que falar, disse com o seu novo ar apolíneo, mas com as mãos crispadas, escondidas nos bolsos: “Não, não é ele”. Custava-lhe que uma só das suas sobrancelhas enfrentasse o cárcere ou que os seus olhos vissem de frente a cadeira eléctrica (sabe-se lá que condenações inspirava a conduta daquele impostor). Saiu e ficou à espera, do outro lado, no café. Queria dizer-lhe “isto não fica assim”, mas não o viu sair e já era escuro quando desistiu. Nessa noite encontrou o suspeito em sonho e aconteceu aquilo que sempre temera: ser reconhecido. Desde então, não há sonho que não lhe ocupe a memória o dia inteiro. Mas, por mais que contemple a fotografia que traz no bolso, não se lembra de quem é. Parece-lhe ter nas mãos um daqueles ditadores fotografados de frente, cujo semblante *de ninguém* conquista a confiança popular, envergando o slogan “dar a cara pela eternidade”.

Elliot Rain [tradução de Emílio Remelhe]

5 Reconstrução de retratos com a BD normalizada

Existem um conjunto de artigos e de trabalhos sobre BDs de rostos humanos em que se realça a importância da normalização dos retratos. Temos de confessar que, em função dos resultados obtidos, inicialmente com a BD faces94, e depois com a nossa BD, tínhamos algum cepticismo relativamente aos resultados a esperar dessa normalização, a menos de uma reconstrução mais rápida, com menos componentes, o que confirmamos com os nossos testes.[17]

Normalizámos os retratos da BD, garantindo o mesmo posicionamento dos olhos em todas as imagens. Para isso os retratos foram sujeitos a operações de rotação, translação, escalamento e corte. Garantimos desta forma que todos os retratos coincidem nas duas pupilas. Nesta secção apresentamos e comparamos os resultados com os que obtivemos com a BD original, para concluir da importância desta normalização.

Como se observa na Figura 112 a média e o desvio padrão têm uma nitidez muito maior no caso da BD normalizada (imagens da esquerda na figura), como seria de esperar, pois os olhos coincidem, quase parecendo corresponder a um retrato real. A média é o retrato que se obtém calculando a média, pixel a pixel, para todos os retratos, o desvio-padrão obtém-se calculando, a partir da média, o desvio padrão, pixel a pixel, para o conjunto dos 439 retratos.

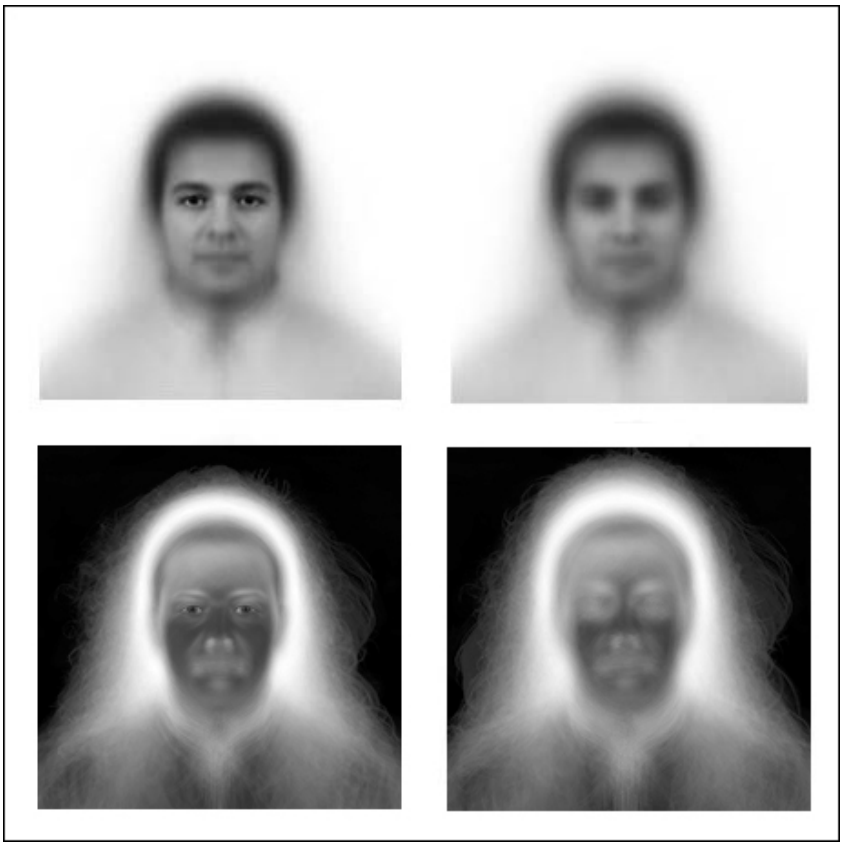


Figura 112 – Imagem média e imagem desvio padrão nas duas BDs

Um outro aspecto interessante poderá ser observar que a média e desvio padrão poderiam ser tomadas por retratos masculinos, o que já acontecia de certa forma com a BD inicial.

Se observarmos os componentes PCA nos dois casos na Figura 113 verificamos que há um efeito

17 Claro que temos de ter em conta que mesmo na BD por normalizar houve um esforço à partida na tomada de imagens para obter retratos aproximadamente na mesma posição. A BD faces94, apesar de maior variação, também só contém retratos frontais, todos do mesmo tipo.

idêntico. Na BD normalizada (linha inferior) as imagens correspondentes às mesmas *eigenfaces* parecem ter maior definição e maior contraste.

Figura 113 – Alguns componentes PCA (K= 5, 10, 30, 50, 100) nas duas BDs

Se observarmos os componentes ICA obtidos para os dois casos – ver Figura 114 – verificamos que há também aqui um efeito idêntico. Na BD normalizada (linha inferior) as imagens correspondentes aos componentes ICA parecem ter maior definição e um maior contraste.

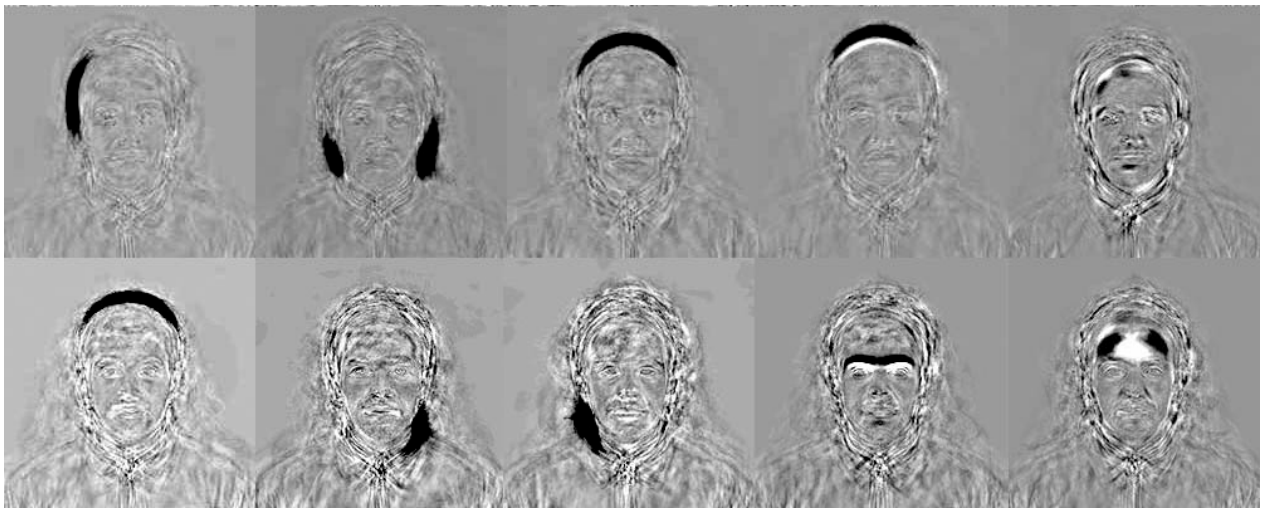


Figura 114 – Alguns componentes ICA (K= 1, 5, 6, 15, 121) nas duas BDs

Se observarmos os componentes NMF (GP com 100 componentes) obtidos para os dois casos – ver Figura 115 – verificamos que há aqui um efeito mais misterioso. Na BD normalizada (linha inferior) as imagens correspondentes aos componentes NMF parecem ter características diferentes. Este resultado deve estar relacionado com o resultado obtido na reconstrução dos retratos, que descrevemos à frente.

Figura 115 – Alguns componentes NMF (1, 2, 3, 4, 5) (GP) nas duas BDs

Nas secções seguintes analisamos os resultados que obtivemos com a BD normalizada (BDN) para cada uma das técnicas estatísticas. Comparamos esses com os resultados anteriores obtidos com a BD inicial (BDI), que apresentámos na secção 4.7. Comparamos os resultados para o conjunto dos retratos da BD e para o conjunto dos retratos fora da BD.

5.1 Comparação entre as duas BDs para todos os retratos da BD

Começámos com o conjunto dos retratos que fazem parte da BD e comparámos os resultados obtidos com a BD normalizada, usando o PCA, ICA, NMF e PCA+ICA, com os resultados anteriores, da secção 4.7.1, calculados com a BD inicial.

5.1.1 Erro na reconstrução para as duas BDs

Na Figura 116 pode observar-se que a evolução do erro calculada para a BD normalizada, com o PCA, é muito parecida com a que se obtivera na BD inicial.

A distribuição dos erros para todos os retratos da BD normalizada (a verde) tem uma menor variância do que na BD por normalizar (a azul) e o erro médio (a vermelho) decresce um pouco mais rapidamente do que na BD inicial (a magenta) mas não se nota qualquer mudança extraordinária, confirmando o que prevíramos. A normalização permite apenas uma convergência um pouco mais rápida, o que era expectável.

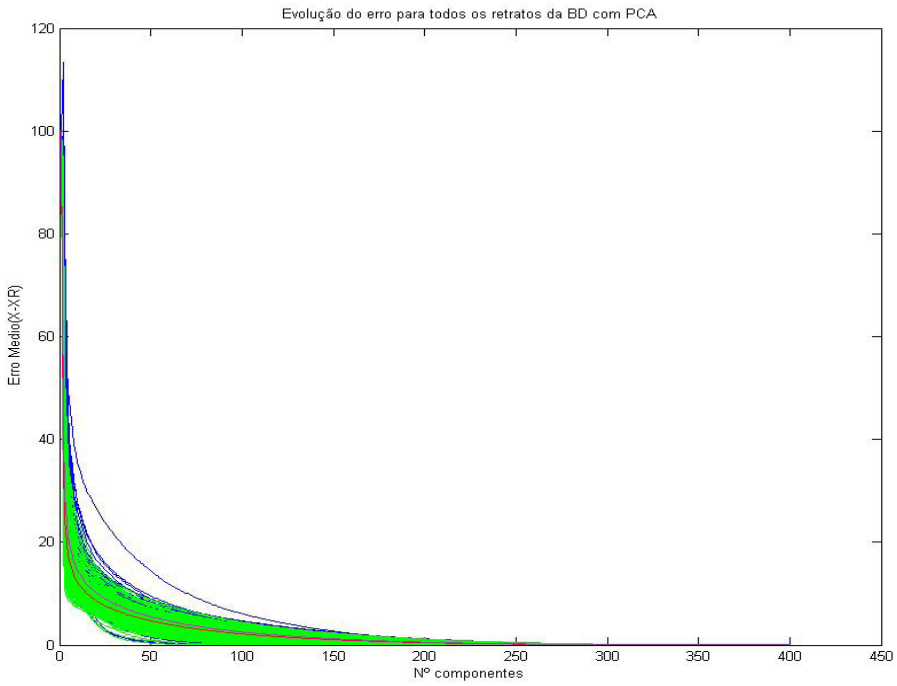


Figura 116 – Evolução do erro no PCA para a BD e a BD normalizada

No caso do ICA, observamos na Figura 117 que, apesar de haver uma menor dispersão das curvas de erro (a verde) do que existia na BD inicial (a azul), a curva com a evolução do erro médio na BD normalizada (a magenta) quase coincide com a que obtivéramos para a BD inicial (a vermelho). Ao contrário do PCA a normalização parece não introduzir qualquer melhoria no ICA.

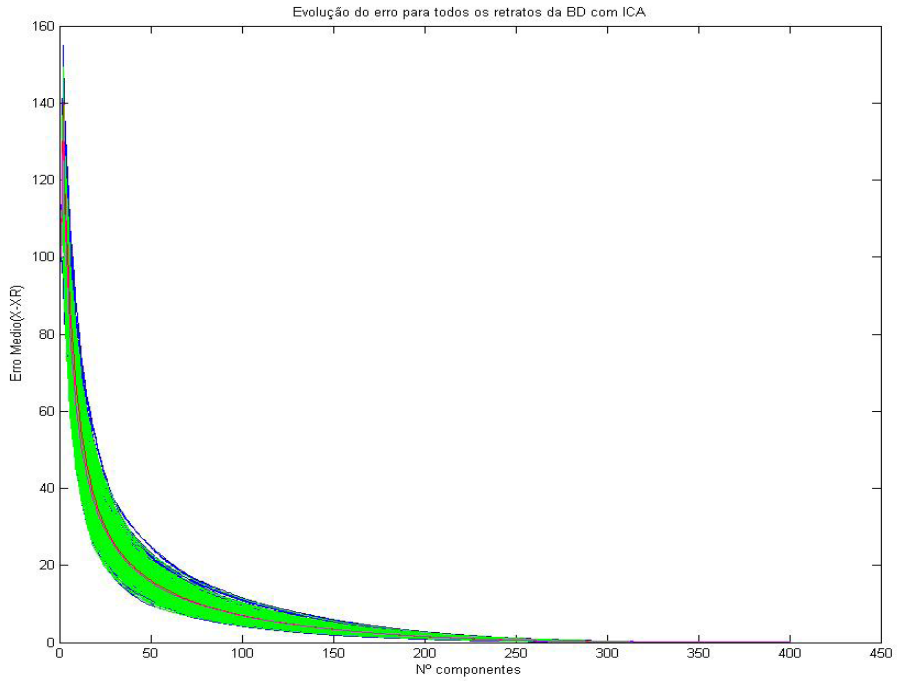


Figura 117 – Evolução do erro no ICA para a BD e a BD normalizada

Tentámos fazer uma avaliação idêntica à que realizámos para o PCA e o ICA com o NMF, utilizando o algoritmo GP com 100 componentes, de acordo com aquilo que tínhamos feito com a BD inicial.

Fomos surpreendidos por resultados que contrariaram aquilo que esperávamos obter. O algoritmo calculou os componentes, parecendo convergir, mas aos calcular o erro para todos os retratos da BD observámos que apenas se conseguia uma reconstrução aceitável – aceitável significa falar de valores próximos dos que obtivéramos para o NMF, à volta de 6 a 7 em 100 – para um pequeno número de retratos, mas para a maioria obtivemos erros elevadíssimos, que revelam que o processo de reconstrução falha. Estes resultados sugerem talvez que estamos em presença de uma solução quase trivial, típica do NMF, mas curiosamente apenas para uma fracção dos retratos da BD. Como este processo tinha revelado alguma instabilidade por exemplo com as inicializações tentámos variar alguns parâmetros como o nº de iterações ou de componentes, mas a situação não se modificou, ou seja o NMF parece ter sido perturbado pela normalização da BD e não conseguimos obter resultados com significado para os 400 retratos! Este resultado mereceria porventura uma investigação que não tivemos oportunidade de aprofundar.

Finalmente no sistema híbrido PCA+ICA observamos na Figura 118 que distribuição dos erros para todos os retratos da BD normalizada, BDN, (a verde) tem uma menor variância do que na BD por normalizar (a azul) e o erro médio (a negro) decresce um pouco mais rapidamente do que na BD inicial (a vermelho) seguindo um padrão idêntico ao do PCA.

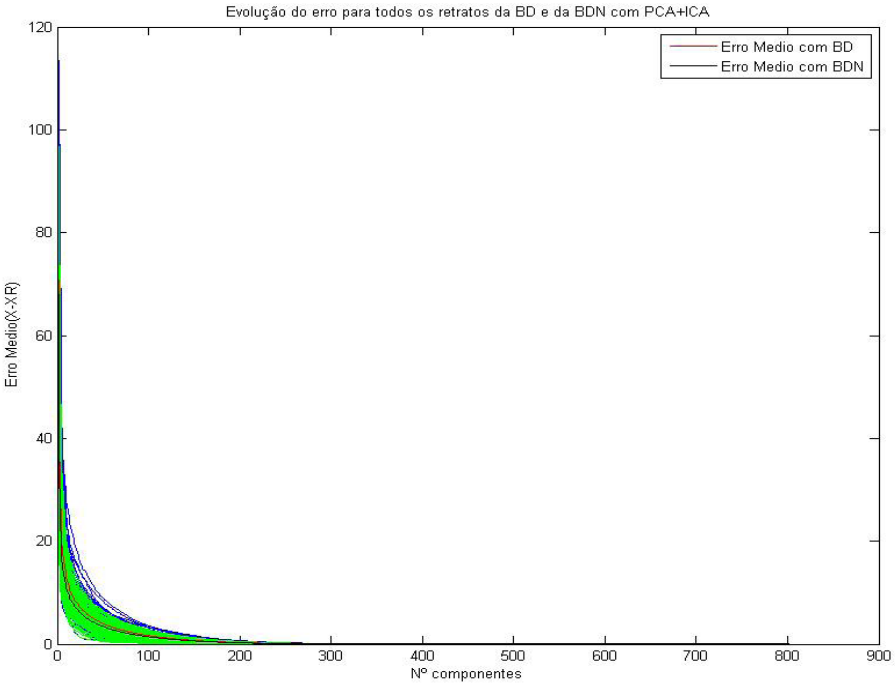


Figura 118 – Evolução do erro no PCA+ICA para a BD e BD normalizada

A normalização permite, uma vez mais apenas, uma convergência um pouco mais rápida, o que era expectável.

Observou-se ainda que o sistema calculado pelo algoritmo *pesqproj* utilizado para decompor os retratos neste sistema não ortogonal apenas utiliza cerca de 600 componentes dos 800 possíveis.

5.1.2 Número componentes necessários para as duas BDs

A Tabela 12 permite confirmar o resultado obtido com o PCA, na secção anterior, para o erro de reconstrução.

BD	Erro<12	Erro<10.5	Erro<9	Erro<7.5	Erro<6	Erro<4.5	Erro<3	Erro<1	Erro<0.1
BDI	K=13	K=18	K=23	K=33	K=44	K=61	K=88	K=161	K=281
BDN	K=9	K=13	K=18	K=25	K=36	K=52	K=77	K=152	K=275

Tabela 12 – Limiares de erro médio PCA para um retrato da BD

Consegue-se reconstruir um retrato impondo um erro < X com um número menor de componentes. Esta diferença é mais significativa para erros elevados. Por exemplo para termos um erro < 12 precisamos de 9 em vez de 13 componentes, ou seja cerca de 31% menos. Para assegurarmos um erro < 4.5, que como veremos assegura um reconhecimento quase certo, precisamos de 52 em vez de 61 componentes, o que corresponde a uma economia de 15% de componentes.

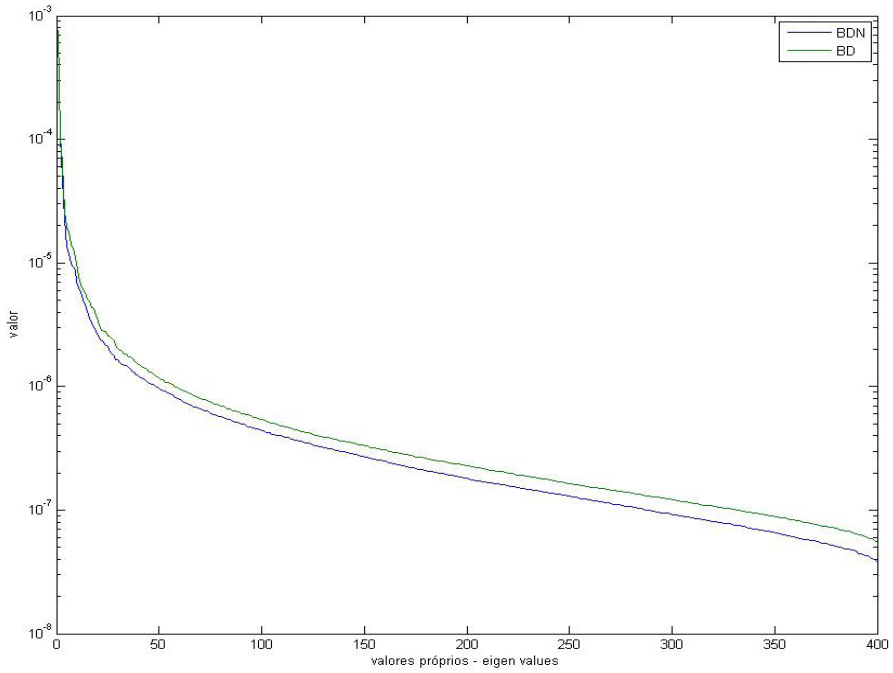


Figura 119 – Valores próprios no PCA

A Figura 119 apresenta as curvas que desenham os valores próprios dos vectores próprios de 1 a 400, calculados no método PCA – usámos uma escala logarítmica para conseguir distinguir melhor as duas curvas – observamos que com a BDN a curva se desloca para a esquerda e que há uma inflexão das curvas próxima do valor 50 ou 60, respectivamente com BDN e BDI. A partir deste valor, os valores próprios decrescem lentamente para zero. Este valor é muitas vezes referenciado como o limite a partir do qual os componentes podem ser ignorados com pequenos erros. Com a BDN este valor diminui, reflectindo a melhoria na reconstrução com o PCA, que precisa de menos componentes. Como se constata com 50 ou 60 componentes obtemos, para BDN e BDI respectivamente, um erro de reconstrução de cerca de 4.5 que, como veremos, assegura um reconhecimento quase certo.

No caso do ICA, a Tabela 13 permite confirmar o resultado obtido para o erro de reconstrução na secção anterior.

BD	Erro<12	Erro<10.5	Erro<9	Erro<7.5	Erro<6	Erro<4.5	Erro<3	Erro<1	Erro<0.1
BDI	K=65	K=73	K=81	K=93	K=110	K=129	K=156	K=223	K=315
BDN	K=64	K=72	K=81	K=93	K=109	K=130	K=158	K=225	K=317

Tabela 13 – Limiares de erro médio ICA para um retrato da BD

Observa-se que no ICA os valores para o número de componentes necessários para garantir erro < X quase coincide com o que obtivéramos com a BD inicial.

Finalmente para o PCA+ICA, tal como no PCA, consegue-se com a normalização da BD reconstruir um retrato, impondo um erro < X, com um número menor de componentes. Na Tabela 14 vemos que esta diferença é mais significativa uma vez mais para erros elevados. Por exemplo para termos um erro < 12 precisamos de 8 em vez de 11 componentes, ou seja cerca de 27% menos. Para assegurarmos um erro < 4.5 que, como veremos na secção 6, assegura um reconhecimento quase certo, precisamos de 37 em vez de 45 componentes, o que corresponde a uma economia de 18% de componentes.

BD	Erro<12	Erro<10.5	Erro<9	Erro<7.5	Erro<6	Erro<4.5	Erro<3	Erro<1	Erro<0.1
BDI	K=11	K=14	K=18	K=24	K=32	K=45	K=67	K=133	K=259
BDN	K=8	K=10	K=14	K=18	K=25	K=37	K=57	K=121	K=252

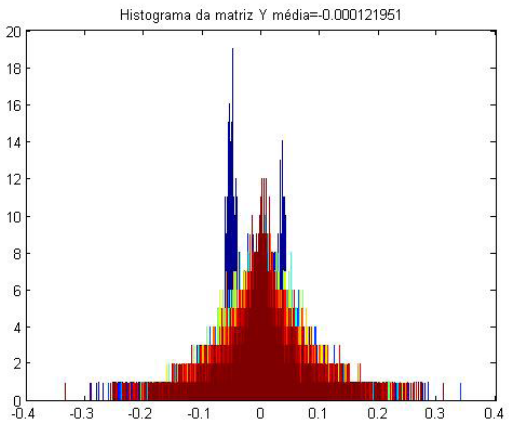
Tabela 14 – Limiares de erro médio PCA+ICA para um retrato da BD

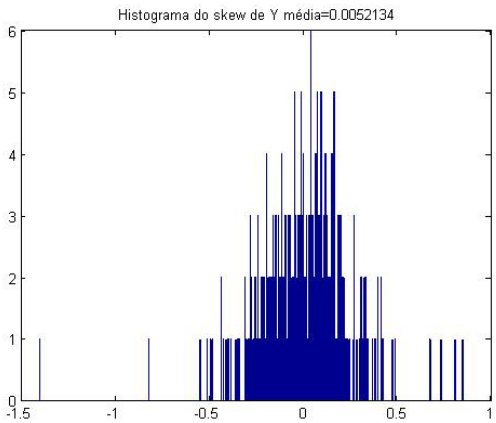
5.1.3 Avaliação das estatísticas dos coeficientes para as duas BDs

As curvas que obtivemos para as estatísticas de 3ª e 4ª ordem no PCA – ver Figura 120 – não diferem substancialmente das que obtivéramos para a BD não normalizada da Figura 99. O *kurtosis* médio em especial, que era 1.11, passa a 1.13.

No caso do ICA as curvas que obtivemos para as estatísticas de 3ª e 4ª ordem – ver Figura 121 – também não diferem substancialmente das que obtivéramos para a BD não normalizada da Figura 99; existe um aumento ligeiro do *kurtosis*. O *kurtosis* médio que era 10.7 é com a BD normalizada de 12.6.

Finalmente para o PCA+ICA, as curvas que obtivemos para as estatísticas de 3ª e 4ª ordem – ver Figura 122 – também quase não diferem das que obtivéramos para a BD não normalizada da Figura 99. O *kurtosis* aumenta muito pouco de 343.4 para 344.8.





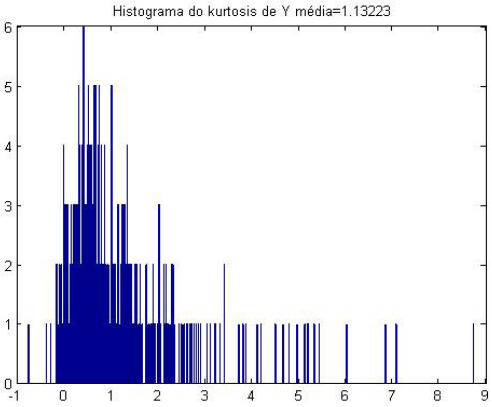
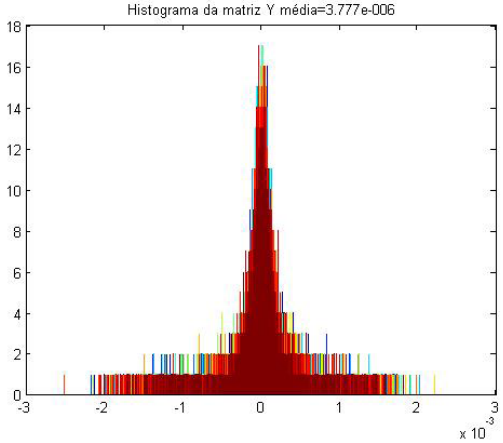
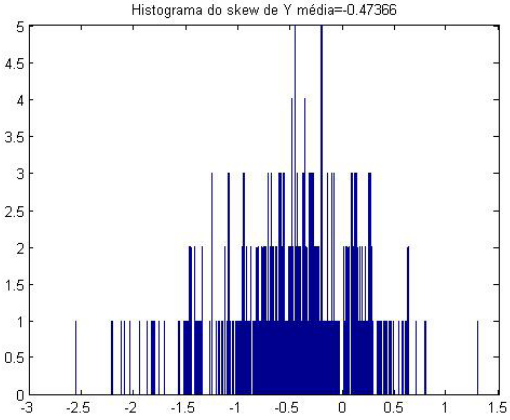


Figura 120 – Estatísticas para a BD normalizada para o PCA





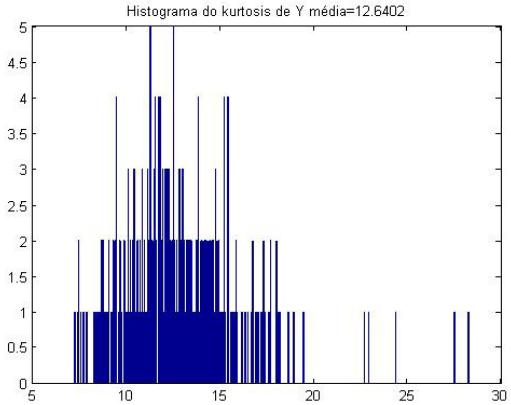


Figura 121 – Estatísticas para a BD normalizada para o ICA

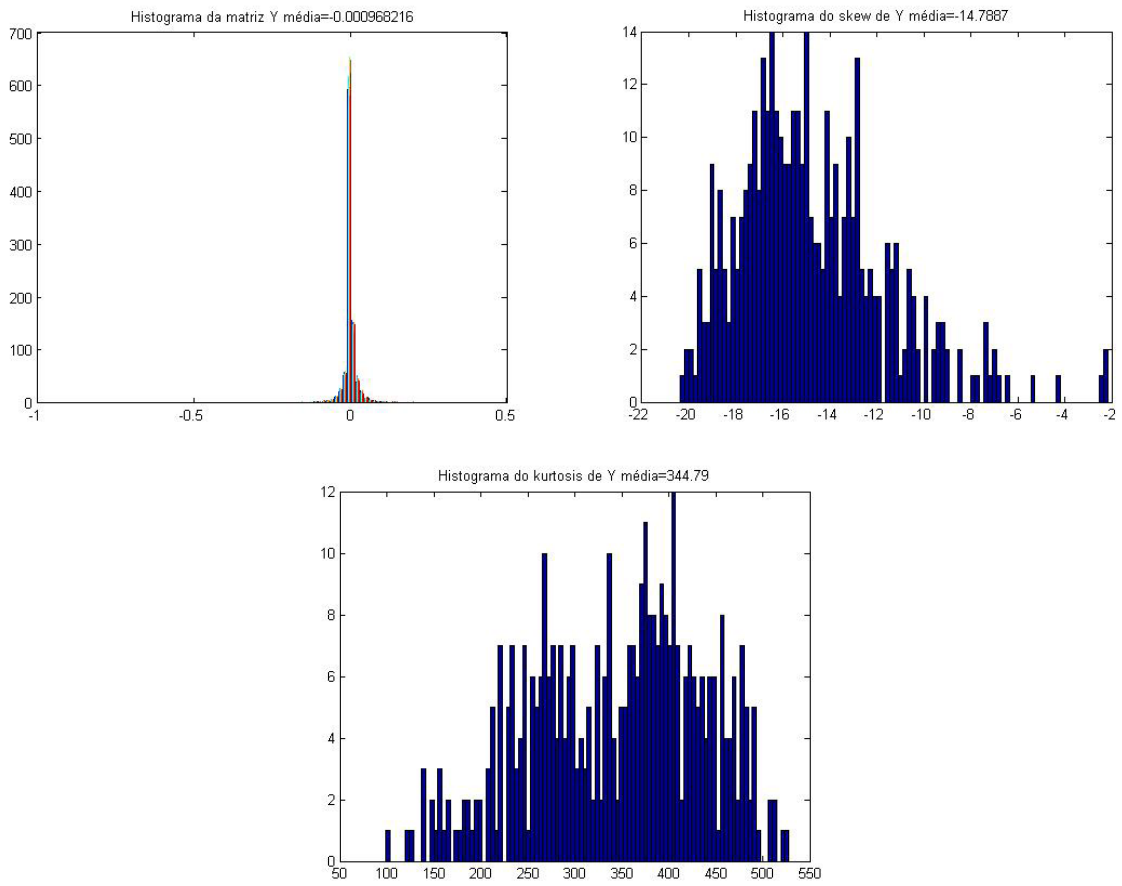


Figura 122 – Estatísticas para a BD normalizada para o PCA+ICA

Estes resultados indiciam que as conclusões que tirámos, quando analisámos os mesmos histogramas para a BD não normalizada, na secção 4.7.1.3, parecem manter-se válidas.

5.2 Comparação entre as duas BDs para todos os retratos fora da BD

Fazemos nesta secção uma comparação idêntica à que fizemos na secção 5.1, para o PCA, ICA e PCA+ICA, mas desta vez para o conjunto dos retratos que não fazem parte da BD. Os resultados são comparados com os resultados obtidos com a BD inicial, apresentados na secção 4.7.2.

5.2.1 Erro na reconstrução para as duas BDs

Mais uma vez as curvas que se obtêm para o erro médio da reconstrução são idênticas às que obtivéramos para a BD inicial. Confirma-se que com a BDN a convergência é mais rápida e o erro final mais baixo.

Observámos que o erro médio final baixa de 6.5 para 5.1, para todas as técnicas estatísticas ou seja, o erro final na reconstrução, com a BD normalizada, baixa apenas ligeiramente. Observámos também que esse erro final varia entre 3.4 e 8.7, para o conjunto das 39 imagens fora da BD, em vez de entre 4.6 e 10.2, que obtivéramos com a BD inicial.

Mais uma vez, ao observarmos apenas os valores finais do erro, as diferenças não parecem ser espectaculares. Observa-se no entanto, como se vê na Figura 123, para o caso do PCA, que as

curvas convergem para o erro final bastante mais rapidamente do que na BD inicial.

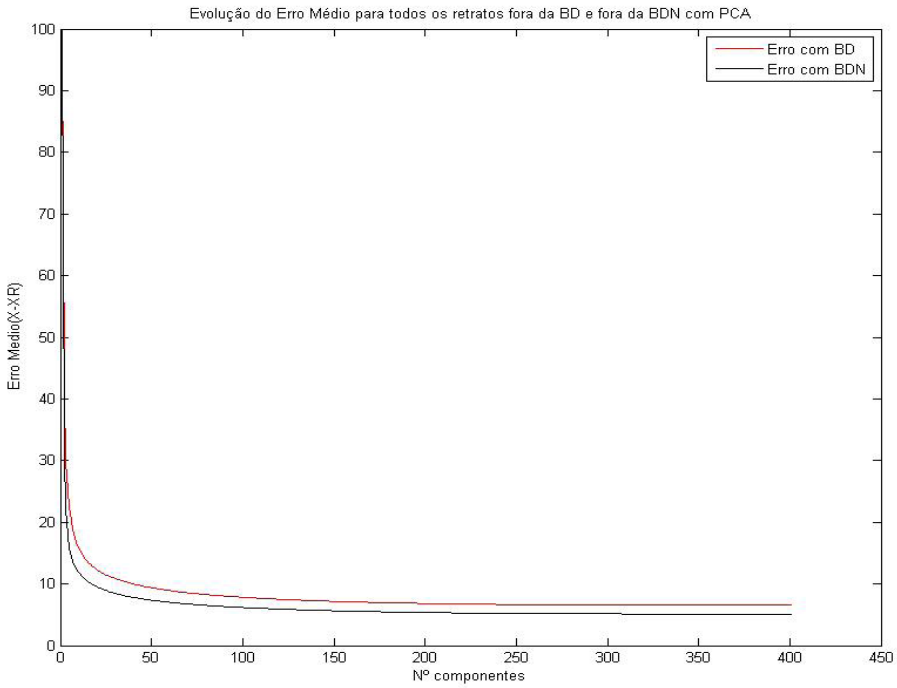


Figura 123 – Evolução do erro médio no PCA para as duas BDs

A Figura 124 mostra a evolução do erro médio para as três técnicas: PCA, ICA e PCA+ICA.

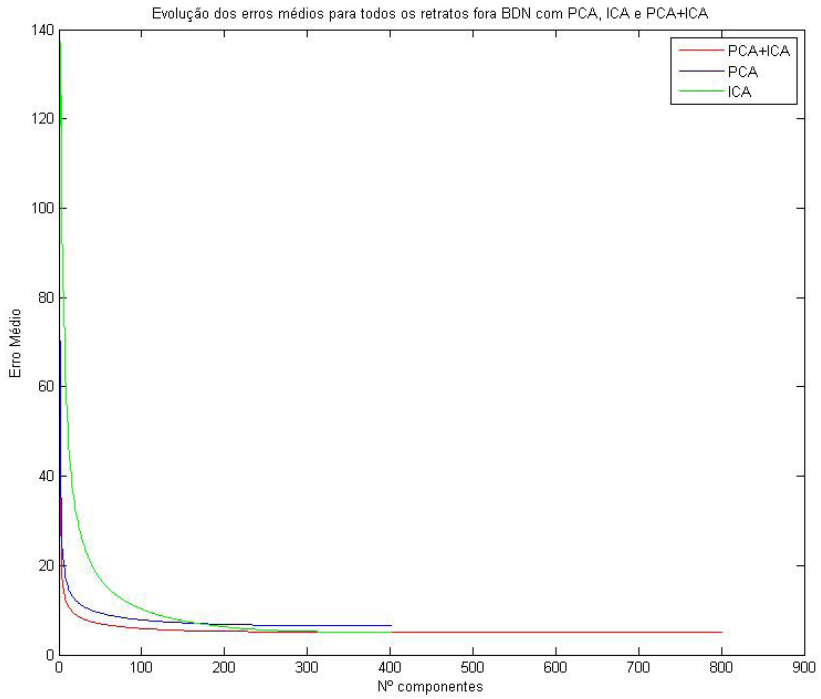


Figura 124 – Evolução do erro no PCA, ICA e PCA+ICA para a BD normalizada

O erro final é o mesmo, 5.1, mas claramente, uma vez mais, o PCA+ICA permite uma reconstrução bastante mais rápida, ainda que o PCA.

5.2.2 Número componentes necessários para as duas BDs

Os resultados da Tabela 15, onde vemos o número de componentes necessários para assegurar erro < X, permitem-nos observar melhor que a melhoria introduzida pela BDN já é bastante mais espectacular, quando olhamos para a rapidez da reconstrução dos retratos fora da BD. Conseguem-se os mesmos erros com muito menos componentes do que com a BD inicial. Por exemplo para um erro < 9 precisamos no PCA de 24, em vez de 57 componentes: menos de metade do valor inicial. Consegue-se mesmo uma curva quase coincidente com a curva de evolução do erro para retratos dentro da BD, para erros > 10, como observamos, se comparamos com os resultados da Tabela 12.

No caso do ICA, apesar de existir uma melhoria efectiva, ela é muito menos evidente do que a observada no PCA.

A melhoria introduzida pela BDN é ainda mais espectacular no que diz respeito à rapidez da reconstrução dos retratos fora da BD com o PCA+ICA. Tal era expectável pois o sistema híbrido junta à melhoria do PCA a melhoria do ICA, tal como já acontecera na BD inicial. Conseguem-se os mesmos erros ainda com menos componentes do que com o PCA. Por exemplo para um erro < 9 precisamos de 20, em vez de 24 componentes do PCA, que já reduzira dos 57 que necessitávamos com a BD inicial: ou seja quase atingimos um terço do valor inicial para o PCA.

Técnica	BD	Erro Final	Erro<12	Erro<10.5	Erro<9	Erro<7.5	Erro<6
PCA+ICA	BDI	6.5	K=18	K=28	K=48	K=100	-
	BDN	5.1	K=8	K=12	K=20	K=38	K=93
PCA	BDI	6.5	K=21	K=33	K=57	K=120	-
	BDN	5.1	K=9	K=14	K=24	K=46	K=112
ICA	BDI	6.5	K=92	K=113	K=145	K=205	-
	BDN	5.1	K=80	K=96	K=119	K=153	K=216

Tabela 15 – Nº comp. necessários (N) para erro < X - retratos fora da BD – com BDI e BDN

Se tivermos em conta também os resultados da secção 6, concluímos que pode fazer-se com PCA+ICA uma reconstrução de um retrato, mesmo se fora da BD, com reconhecimento positivo efectivo, com apenas cerca de 20 componentes, em vez dos 48 que obtivéramos.

Parece ser possível reconstruir todos, mas mesmo todos os retratos, retratos de toda a gente, com apenas cerca de 20 componentes, garantindo um erro < 9, o que de acordo com os resultados da secção 6 possibilita um reconhecimento efectivo.

5.3 Erro fotográfico para as duas BDs

Comparámos as imagens reconstruídas com a BD normalizada e as imagens reconstruídas com a BD inicial, utilizando as imagens reconstruídas com o PCA e com o ICA. Como se observa na Figura 125 parece-nos que continua a ser possível associar um valor do erro geométrico com a qualidade fotográfica do retrato, ainda que na BD normalizada haja claramente uma melhor definição para erros mais altos, sobretudo na zona dos olhos.

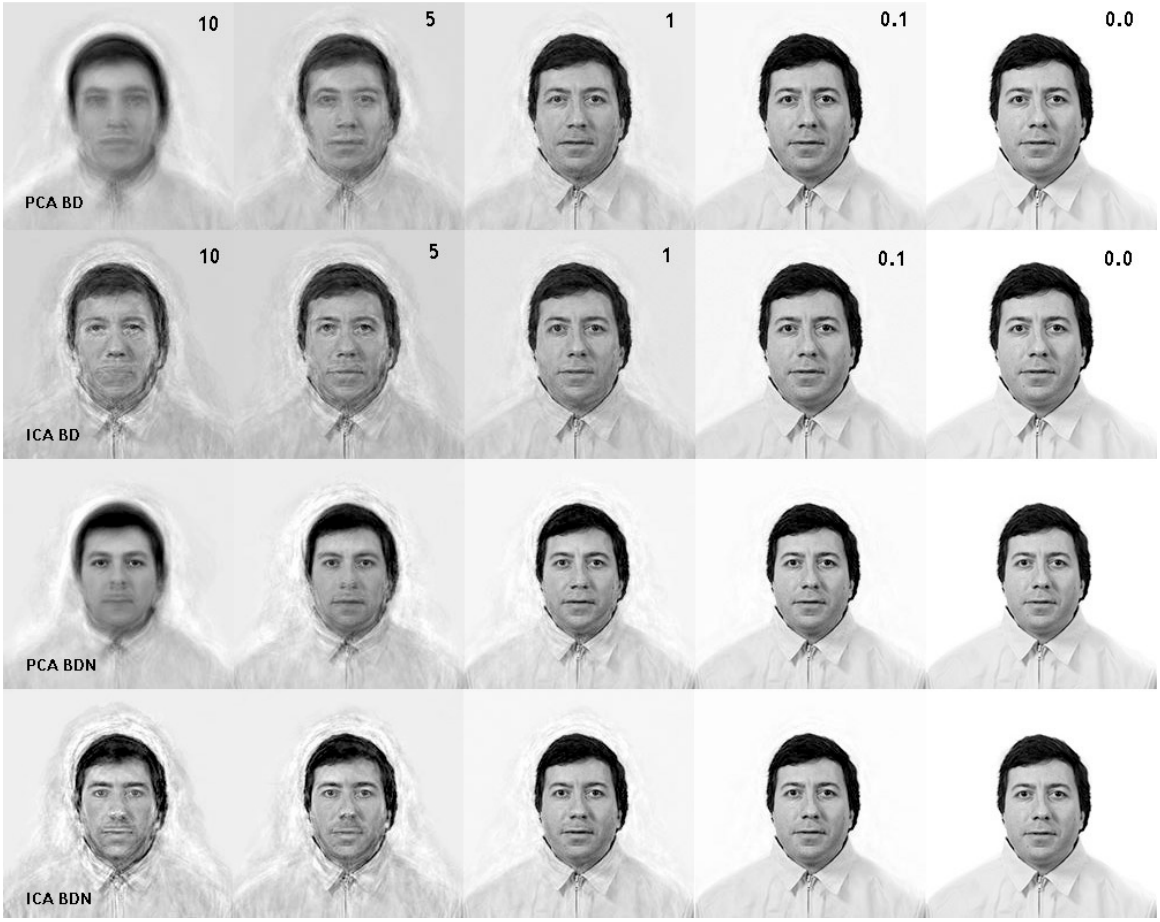


Figura 125 – Qualidade fotográfica para a BD e a BD normalizada

Na Figura 125 temos o retrato reconstruído pelo PCA e pelo ICA, com erros de 10, 5, 1 e 0.1, com a BD inicial e com a BDN (BD Normalizada). Mais uma vez o ICA parece produzir imagens um pouco mais definidas para erros elevados, ver erro=10.

É também interessante comparar os resultados para imagens que não pertençam à BD. Para a comparação ser mais fácil decidimos repetir todos os cálculos com outra BDN, onde a imagem que temos utilizado para testes não foi integrada, e depois reconstruímo-la a partir dos novos componentes PCA obtidos. A Figura 126 mostra os retratos reconstruídos com PCA quando o retrato fazia parte da BD (1ª linha) e quando era um retrato inteiramente novo. No segundo caso o erro final é e 5.7. Verificamos uma vez mais que as imagens com o mesmo erro geométrico parecem idênticas.

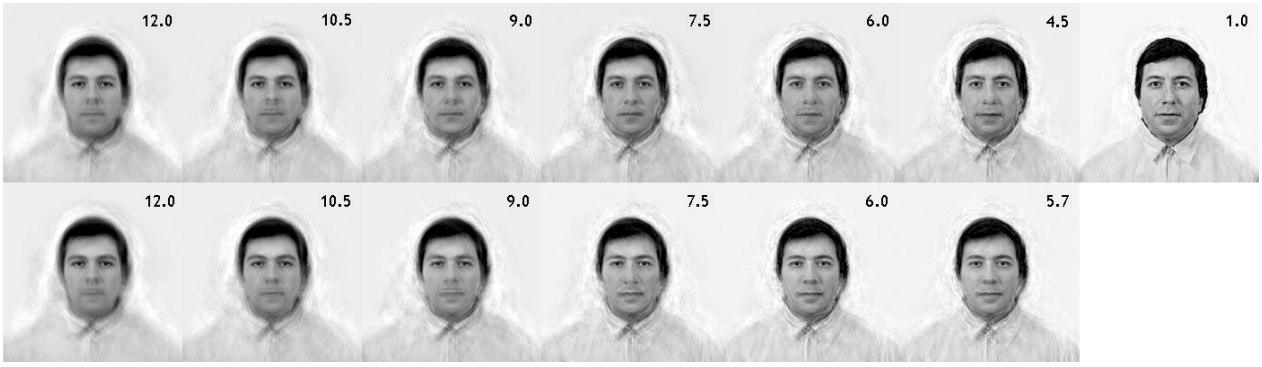


Figura 126 – Qualidade fotográfica para uma imagem nova – BD e BD normalizada

Este resultado, associado ao resultado nos inquéritos apresentados na secção 6, permite-nos reafirmar que o sistema possibilita não só reconstruir sem erro todos os retratos da BD, mas

reconstruir todos os novos retratos com um erro que garante um reconhecimento quase perfeito, pois o erro final é pouco superior a 4.5 (Como vimos obtivemos um erro médio=5.1).

Resultados idênticos poderiam ser mostrados com o ICA ou PCA+ICA.

Este resultado parece ser muito relevante para o conceito de "*unsupervised learning*" de novos retratos e para o conceito de "*sparse coding*": **verificamos que com 400 componentes, calculados a partir de 400 retratos (200x200 px) de uma BD, conseguimos reconstruir com erro, mas com um valor de erro que permite o reconhecimento quase perfeito de todos os retratos, inclusive novos retratos que não foram utilizados no cálculo dos componentes.** Os componentes permitem "aprender" novos retratos sem recalcular os componentes. Em rigor para obter erros com valores próximos de 6, por exemplo, são necessários apenas cerca de 25 ou ou 90 componentes, respectivamente para retratos da BD ou fora da BD. De realçar que mesmo com a BD não normalizada se poderia ter tirado uma conclusão idêntica. Apenas o erro final era maior (cerca de 6.5). Obtemos nos dois casos, de acordo com os resultados da secção 6, valores na zona crítica para o reconhecimento.

Na secção 7 pretendemos ir mais longe nesta análise, à custa da avaliação de algoritmos para integração de duas BDs.

5.4 Conclusões sobre o uso da BD normalizada

Ao compararmos os resultados que obtivemos com a BD normalizada – fazendo coincidir os dois olhos em todos os retratos - com os resultado que tínhamos obtido com a BD inicial, sem qualquer normalização verificamos que

- As imagens que correspondem à média e ao desvio padrão adquirem uma vivacidade, fruto sobretudo da grande definição da zona dos olhos
- A relação erro geométrico/erro fotográfico não parece ser muito diferente da que obtivéramos com a BD por normalizar, confirmando existir essa relação
- As curvas de evolução do erro de reconstrução dos retratos da BD normalizada têm um andamento idêntico às que obtivéramos para a BD por normalizar
- No PCA e no PCA+ICA a reconstrução é um pouco mais rápida com a BD normalizada; consegue-se reconstruir um retrato com menos componentes ainda que a diferença não seja espectacular
- No entanto a melhoria na velocidade de reconstrução (*sparse coding*) é muito mais acentuada para retratos fora da BD, onde se consegue o mesmo erro de reconstrução com menos de 50% dos componentes
- O erro final na reconstrução de retratos fora da BD também diminui, ainda que essa diminuição não seja muito grande
- No ICA, as curvas da evolução do erro são quase coincidentes para as duas BDs
- Os valores do *skewness* e de *kurtosis* são idênticos para as duas BDs, as diferenças parecem ser residuais
- A não convergência no NMF com a BD normalizada – que nos surpreendeu – é mais um dado a somar a todos os resultamos, em certa medida decepcionantes, obtidos com o NMF

Em resumo, para os retratos da BD, a menos da espectacularidade das médias e desvio padrão, parece não existir, fruto da normalização da BD, qualquer mudança espectacular, para os retratos da BD, mas apenas uma reconstrução ligeiramente mais rápida nos sistemas que envolvem o PCA (ver nota 17). Tal é particularmente visível para novos retratos onde, para além do erro ser ligeiramente menor, se consegue uma reconstrução muito mais rápida, com um número de componentes K entre metade a um terço do valor inicial.

A face humana nunca mente. É o único mapa que regista todos os territórios onde vivemos.

Luís Sepúlveda

AFINAL, A MEMÓRIA

Afinal eram iguais os homens
as mulheres
vistos de cima
quando abanavam ligeiramente a cabeça
para a frente e para trás
ao mesmo tempo,
ou se inclinavam nas horas da infância, da minha infância,
ou quando mexiam no cabelo uns dos outros
para eu adormecer.
Afinal a memória era um lugar parecido
com a memória, e o sonho era um lugar parecido
com a memória, e nós talvez fôssemos todos, na verdade,
parecidos
uns com os outros.

Filipa Leal, em ***O Problema de Ser Norte***

6 Análise dos resultados dos inquéritos

Os inquéritos foram preparados, utilizando como alvo os alunos e alguns professores do Mestrado Multimédia: um universo de cerca de 50 pessoas. O inquérito tinha como objectivo avaliar a qualidade subjectiva dos retratos reconstruídos com poucos componentes, através do reconhecimento e tentar responder às seguinte questões:

- Qual o erro necessário (e consequentemente o nº de componentes) para possibilitar uma identificação?
- Qual a discriminação que se consegue à medida que se faz diminuir o erro?

A resposta a estas perguntas permitiu-nos avaliar, do ponto de vista do reconhecimento, os resultados conseguidos nas secções 4 e 5, para a reconstrução de retratos de face humana, utilizando as várias técnicas estatísticas. Deu-nos também dados importantes para a avaliação dos algoritmos da secção 7**Error! Reference source not found.**.

6.1 Organização do inquérito

O inquérito contempla 24 retratos de pessoas conhecidas do mestrado e 24 retratos de pessoas desconhecidas fotografadas no ISEP e na FEUP; utiliza 4 imagens construídas com cada um de 6 limiares de erro diferentes (24=4x6). Esses 6 limiares para o erro são: erro = 12.0, 10.5, 9.0, 7.5, 6.0 e 4.5. Estes valores foram escolhidos por forma cobrir a zona crítica para o reconhecimento de um retrato. O número de imagens igual a 4, para cada limiar de erro, não foi fixado arbitrariamente: como se dispunha de retratos de 24 pessoas do mestrado e como se decidiu avaliar o reconhecimento para 6 níveis de sinal, cobrindo a zona crítica suspeita para o reconhecimento de um retrato, optou-se por usar 4 retratos de pessoas diferentes para cada nível de sinal, evitando assim ter de reutilizar um retrato da mesma pessoa com erros diferentes, o que poderia contaminar as respostas. De facto, se por exemplo utilizássemos no inquérito retratos da mesma pessoa com erros de 10.5 e 4.5, como os dois que a Figura 127 mostra, parece-nos que poderia ser possível haver quem identificasse o retrato com erro 10.5 a partir da analogia com a mancha do retrato com erro 4.5. Ou seja, poder-se-ia questionar se também se fariam todas as identificações, sem a ajuda dos retratos da mesma pessoa mas com erro menor.

Figura 127 – Dois retratos da mesma pessoa com erro de 10.5 e 4.5

Para evitar essa possível contaminação optou-se por usar apenas 4 retratos para cada um dos níveis e não os repetir. Claro que esta decisão colocou-nos perante o constrangimento de ter

um pequeno número de retratos por cada nível de erro. Mas como esperávamos receber mais de 25 respostas ao inquérito, ou seja cerca de 50% do universo inquirido, garantíamos mais de 100 respostas para cada nível do sinal, o que nos pareceu suficiente para a análise que pretendíamos realizar.

As 48 imagens utilizadas, distribuídas pelos 6 níveis de erro, foram aleatoriamente distribuídas por 6 páginas de inquérito. Apresentamos o inquérito que realizámos no anexo 10.4.

Figura 128 – Disposição dos retratos por página nos inquéritos

Para cada foto o respondente tentou fazer a identificação, respondendo às seguintes questões:

	Foto1	Foto2	Foto3	Foto4	Foto5	Foto6	Foto7	Foto8
Reconhece a pessoa? (S/N)								
Se reconhece, quem é? (escreva o nome da pessoa)								
Se reconhece, qual o grau de certeza no reconhecimento (Absoluto (A)/ Razoável (R)/ Pequeno (P))								

Tabela 16 – Respostas por retrato no inquérito

A Tabela 17 define a classificação das várias respostas de acordo com a Teoria de Detecção do Sinal (SDT). O inquérito é apresentado em anexo. As secções seguintes apresentam os resultados obtidos e a sua análise.

	ACERTO	**FALHA**	**FALSO ALARME**	**REJEIÇÃO CORRECTA**
Conhece o retratado e identificou correctamente o retrato	X			
Conhece o retratado e identificou erradamente o retrato		X		
Conhece o retratado e não identificou o retrato		X		
Não conhece o retratado e não identificou o retrato				X
Não conhece o retratado mas identificou o retrato			X	

Tabela 17 – Classificações dos tipos de resposta no inquérito

6.2 Resultados do inquérito

Recebemos no total 28 respostas. As 2 tabelas seguintes sintetizam os resultados obtidos. A primeira tabela dá-nos os números de identificações conseguidas para retratos de pessoas do mestrado com erros diferentes, podendo esses reconhecimentos ter sido feitos com grau de confiança pequeno, razoável ou absoluto. A segunda tabela dá-nos os resultados obtidos para retratos de desconhecidos, com todos os valores do erro de reconstrução. Estes últimos retratos tinham como objectivo verificar a possibilidade de falsas identificações, com diversos valores para o erro de reconstrução.

	Não reconhecidas	**Reconhecidas com confiança pequena**	**Reconhecidas com confiança razoável**	**Reconhecidas com confiança Absoluta**	**Total**
Erro=12.0	102	6	4	0	112
Erro=10.5	96	7	8	1	112
Erro=9.0	56	14	26	16	112
Erro=7.5	37	20	25	30	112
Erro=6.0	34	14	14	50	112
Erro=4.5	2	5	13	92	112

	Não reconhecidas	**Reconhecidas com confiança pequena**	**Reconhecidas com confiança razoável**	**Reconhecidas com confiança Absoluta**	**Total**
Retratos de estranhos	664	3	3	2	672

Tabela 18 – Resultados do inquérito

Observa-se facilmente que um erro de 4.5 em 100 na reconstrução de um retrato nos garante quase a 100% o reconhecimento e portanto uma identificação correcta. Num universo de 108 respostas só em dois casos não houve reconhecimento, o que poderíamos se calhar atribuir a distracção, ou porventura a alguma pressa no preenchimento dos inquéritos, pois verifica-se que não só quase todas as pessoas que respondem fazem o reconhecimento correcto nesse casos, mas que quase sempre o fazem com uma certeza absoluta. A certeza absoluta nos retratos com erro=4.5 corresponde a cerca de 82% dos casos, a certeza razoável a apenas cerca de 12% e a certeza pequena a menos de 5%.

Observa-se por outro lado que com um erro de 12% quase não existem reconhecimentos, ou seja o grau do erro parece ser demasiado grande para permitir reconhecer a pessoa no retrato. Apenas em 8% dos casos o reconhecimento se fez. Observa-se que um erro de 10.5 não introduz ainda grande melhoria. Apenas a partir do erro de 9 em 100 se observa uma melhoria efectiva. Com um erro igual a 9 já há uma percentagem de cerca de 50% nos reconhecimentos. Diminuindo o erro para 7,5 e 6 o sucesso aumenta para 66% e 69% respectivamente, para se conseguir praticamente 100%, como vimos, com um erro de 4.5. Verificamos portanto que não se observa uma melhoria muito grande quando o erro evolui de 7.5 para 6, valores que parecem corresponder a uma zona crítica no reconhecimento.

Observa-se por outro lado que quase não há falsos reconhecimentos, aquilo que na SDT se designa por Falsos Alarmes. De facto, em 648 testes, temos apenas 8 falsos alarmes. Este resultado é compreensível pois a complexidade que um reconhecimento de uma face envolve quase não permite identificar um conhecido com uma imagem de outra pessoa, por mais distorcida que esta possa estar, a menos que tivéssemos procurado um grupo de pessoas muito parecidas. Os 8 casos corresponderão muito provavelmente a quase um acto de adivinhação por parte de respondentes, mais inclinados a arriscar uma identificação, do que a um verdadeiro reconhecimento e por isso serão actos tão isolados.

6.3 Avaliação dos resultados do inquérito com a SDT

Utilizámos o modelo de *Signal Detection* Theory – SDT – para avaliar os nossos resultados. Para essa análise recorremos ao SW de cálculo de parâmetros estatísticos chamado *Rscore_Plus*, que é disponibilizado pelo Departamento de Psicologia e de Neurociências da Universidade do Colorado. As curvas que apresentamos e todos os parâmetros estatísticos que caracterizam o modelo SDT foram calculados com esse SW.

O resultado obtido e apresentado para os retratos de estranhos, o pretenso ruído, introduziu uma dificuldade no modelo SDT, pois temos um ruído sem falsos alarmes. Por outro lado, o próprio conceito de ruído, associado a retratos de desconhecidos, parece ter uma sustentação discutível[18]. O modelo SDT pode no entanto ser generalizado para comparar apenas sinais e a capacidade de discriminação entre eles. Decidimos por isso utilizar o modelo para comparar apenas os 6 níveis de sinal. Aliás era esse o objectivo da nossa análise: avaliar como evolui o reconhecimento com a diminuição do erro nos retratos.

Utilizámos um modelo gaussiano, com seis níveis de sinal e com três níveis de critério de decisão, de acordo com o inquérito que descrevemos. Os resultados foram analisados globalmente, sem nos preocuparmos com as respostas de cada uma das pessoas que respondeu ao inquérito. Por um lado não estávamos interessadas em caracterizar cada um dos respondentes, pretendíamos apenas caracterizar um comportamento médio, por outro lado, como vimos, para cada respondente dispúnhamos apenas de 4 retratos por nível do sinal, o que parecia insuficiente para uma análise. Os resultados que obtivemos demonstram a validade do método e o seu significado estatístico.

Depois de experimentar várias soluções alternativas, optámos por permitir desvios padrões diferentes da unidade para os sinais, porque com essa liberdade os resultados estimados eram bastante mais próximos dos observados: o valor de Chi-Square, que de alguma forma mede a validade do modelo, provava claramente a melhor aproximação conseguida.

As duas tabelas sintetizam os resultados obtidos. A primeira apresenta as médias e os desvios padrões das curvas gaussianas para cada um dos sinais e a segunda mostra os níveis de decisão dos observadores, de acordo com o modelo SDT.

	Média	**Desvio Padrão**
S12.0	0	1
S10.5	0.0644	1.1937
S9.0	1.3426	1.2836
S7.5	1.8629	1.2642
S6.0	2.3552	2.0718
S4.5	3.7788	1.2289

	Critério Reconhecimento	**Critério Reconhecimento Razoável**	**Critério Reconhecimento Absoluto**
Xc	1.3331	1.8501	2.6634

Tabela 19 – Parâmetros característicos do modelo SDT

O SW Rscore_plus permite obter todos os parâmetros estatísticos interessantes associados ao modelo SDT e toda a informação necessária para traçar todas as curvas que caracterizam o modelo. A Figura 129 representa as 6 distribuições obtidas, com os parâmetros apresentados na Tabela 19.

18 De facto no nosso trabalho começámos por calcular o modelo SDT, considerando os testes com retratos de desconhecidos como ruído, e o modelo funcionou e teve mesmo um valor de chi-square bastante bom, mas decidimos apesar disso apenas avaliar os 6 sinais, por nos parecer uma abordagem mais consistente.

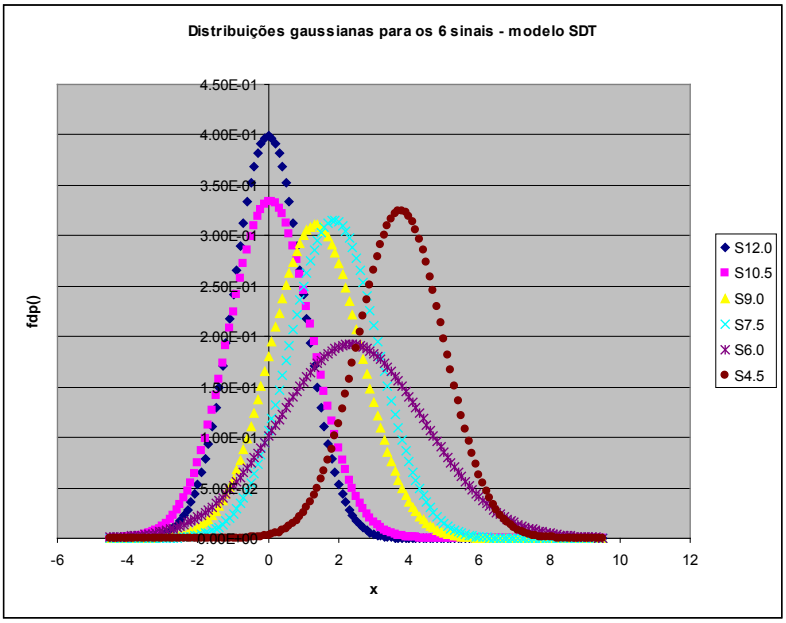
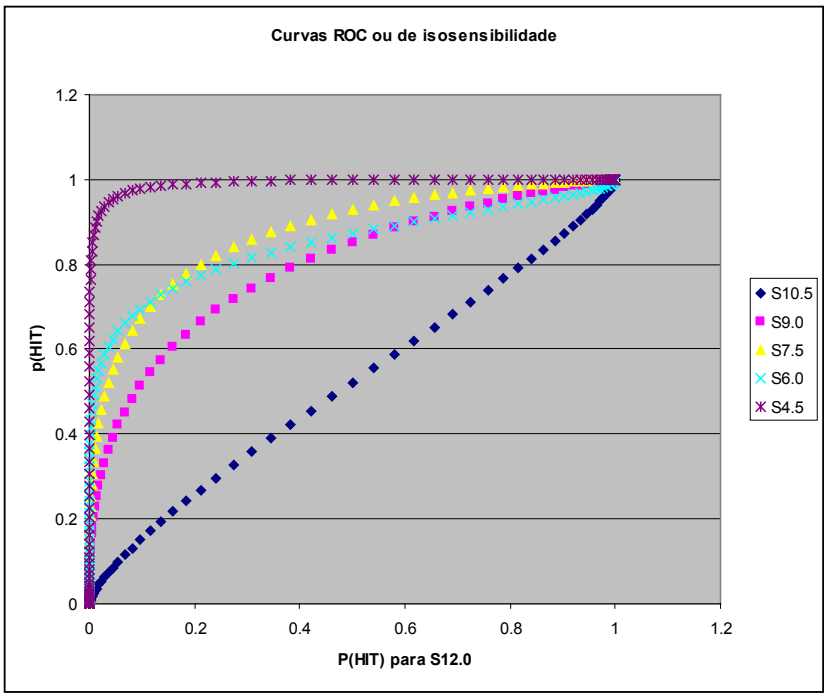


Figura 129 – Modelo gaussiano obtido com SDT

Observa-se que a curva da distribuição para o sinal S10.5 (erro=10.5) quase se sobrepõe à curva para o sinal S12.0 (erro=12.0), com um desvio padrão um pouco maior. Este resultado traduz o facto observado de o erro=10.5 quase não melhorar a probabilidade de reconhecimento. Observa-se que nesta zona são mais determinantes para o reconhecimento características especificas de cada retrato do que o valor do erro. De facto, observa-se que há retratos que são poucas vezes identificados por algumas pessoas mas há outros que nunca são reconhecidos. Para as curvas de S9, S7.5 e S6 observa-se um claro deslocamento para a direita, mas o grande deslocamento observa-se para S4.5. Este resultado traduz o facto observado de ser perto deste valor que o grande salto para o reconhecimento efectivo acontece. Observa-se também que mesmo para os sinais de S9, S7.5 e S6 existem ainda situações de retratos reconhecidos por todos e retratos reconhecidos por poucos. Este resultado não será de estranhar, pois todos sabemos que há rostos que memorizamos e reconhecemos facilmente, que identificamos num olhar, e outros em que tal não acontece.



7

Figura 130 – As curvas ROC ou de isosensibilidade

As curvas traduzem o comportamento de cada sinal, tomando como referência S12.0

Observando as curvas ROC na Figura 130 constata-se que a curva para S10.5 quase coincide com a diagonal, o que traduz comportamentos idênticos para S12 e S10.5. As curvas afastam-se da diagonal à medida que o erro diminui até quase coincidir com a situação perfeita para S4.5.

A Tabela 20 apresenta os valores A(z), ou seja o valor da área limitada pelas curvas, uma das formas mais utilizadas para medir a discriminação entre sinais. Valores próximos de 0.5 indicam pequena discriminação, valores perto de 1 correspondem a uma grande discriminação.

Sinais entre os quais a discriminação é calculada	A(z)
s12.0-s10.5	0.5165
s12.0-s9.0	0.7954
s12.0-s7.5	0.8761
s12.0-s6.0	0.8470
s12.0-s4.5	0.9915

Tabela 20 – Discriminação entre sinais

Observa-se por exemplo que S6 tem um valor de A(z) ligeiramente inferior a S7.5, apesar de a curva de S6 estar situada claramente à direita na Figura 129. Este resultado deriva da maior variância encontrada para S6, o que provoca uma inversão visível na curva, característica de valores de $\sigma > 1$. Acreditamos que esta maior variância pode estar relacionada com o facto de haver em S6 retratos sempre reconhecidos, mesmo com certeza absoluta, e retratos poucas vezes reconhecidos por todos os respondentes.

Parece-nos portanto que as conclusões a tirar poderão ser:

- para um erro entre 12 e 10.5 o reconhecimento é difícil. Não existe quase discriminação entre S12 e s10.5
- para um erro = 9 já existe uma clara discriminação do sinal (quase 0.8) e passamos a reconhecer com alguma facilidade – taxas da ordem de 50%
- para um erro = 7.5 ou 6 o reconhecimento melhora, mas sem grandes diferenças entre eles – a discriminação do sinal relativamente a s12 pouco aumenta
- finalmente para um erro de 4,5 o reconhecimento é quase perfeito. A discriminação do sinal é quase total. O valor de A(z) é quase 1.

(Abdi 1987, 2007); (Harvey 2009); (Heeger 2003); (Wickens 2001)

6.3.1 Validade do modelo

O modelo parece adaptar-se muito bem aos dados do questionário. O modelo tem 15 graus de liberdade. O valor da estatística de Pearson calculado, utilizando a expressão (57) é de 6.725.

$$X^2 = \sum_{i=1}^{c} (x_i - m_i)^2 / m_i \qquad (57)$$

Nesta expressão x_i representa os valores estimados pelo modelo e m_i corresponde aos valores observados. Recorrendo à distribuição chi-square que caracteriza esta estatística, obtém-se uma probabilidade 0.96484, para o valor ser maior do 6.275, ou seja uma probabilidade muito próxima de 1, o que garante a validade do modelo.

A Figura 132 compara os valores estimados pelo modelo comparados com os valores observados no teste e observa-se que se dispõem ao longo da recta onde X=Y, ou seja a estimativa do modelo é muito próxima dos valores observados.

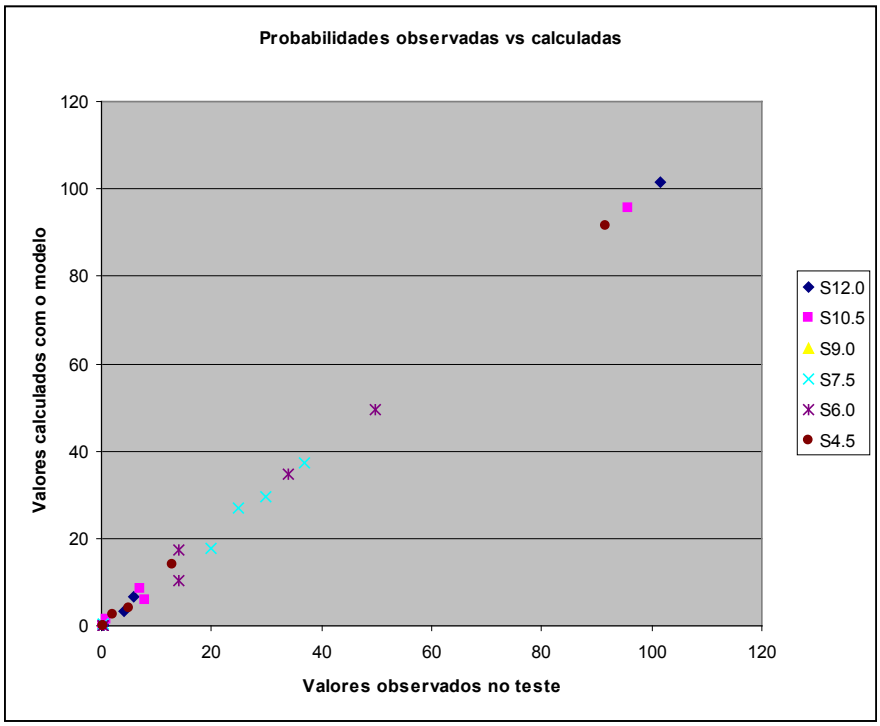


Figura 132 – Valores observados vs valores estimados pelo modelo

(Abdi 1987, 2007); (Harvey 2009); (Heeger 2003); (Wickens 2001)

Na memória ficaram os sinais dos bosques ceifados,
as dunas desfeitas e algumas casas abandonadas

Al Berto

olho-te de dentro

de outra forma
não sei olhar

Luís Tobias

7 Algoritmos para a integração duas BDs de retratos

Se tivermos uma BD de N retratos e se com base nessa BD tivermos calculados os seus componentes PCA ou ICA, como integrar nesse conhecimento novos retratos? A solução óbvia para integrar K retratos novos seria considerar uma nova base de dados com N+K retratos e recalcular todos os componentes. Pretendemos nesta secção validar um algoritmo que permita fazer esta integração com base nos componentes das duas BDs: os N retratos existentes e os K retratos a integrar. Pretendemos em particular limitar o número total de componentes, procurando optimizar o valor do erro de reconstrução para todos os retratos.

Este trabalho foi inspirado por um modelo especulativo de memória, organizada em dois níveis: uma memória de longo e outra de curto prazo, descrito na secção 7.1. Esse modelo deveria assegurar mecanismos para possibilitar a reconstrução de novas imagens da memória de curto prazo na memória de longo prazo. Esse modelo deveria possibilitar a aprendizagem não supervisionada de novos retratos, ou seja de novas imagens de faces.

Apesar de os algoritmos que vão ser descritos, e cujos resultados são apresentados, serem genéricos para o problema apresentado de integrar dois grupos de componentes para dois grupos de retratos, em particular sendo os componentes de tipo diferente, por exemplo PCA e ICA, vamos sempre referirmo-nos, ao designar cada um dos grupos, à memória de longo prazo e à memória de curto prazo, apenas como uma forma de visualizar melhor o problema.

Vamos avaliar e testar dois algoritmos distintos. Vamos começar por avaliar o erro na reconstrução de novas imagens, para um algoritmo mais simples (a que chamaremos algoritmo inicial), onde os componentes da memória de curto prazo são movidos um a um para a memória de longo prazo, possibilitando assim melhorar a reconstrução das novas imagens, obtendo-se um erro mais baixo. Este algoritmo contém no entanto o constrangimento de impor um crescimento gradual do número de componentes na memória de longo prazo. Ora uma memória vai, no contexto real, impor limites físicos. Decidimos, por isso, testar também outro algoritmo, onde um novo componente da memória de curto prazo, em vez de se adicionar, substitui um componente anterior na memória de longo prazo, impondo um número total de componentes constante.

Na avaliação destes dois algoritmos iremos ter em conta os resultados da secção 6, que demonstrou ser possível fazer um reconhecimento de um retrato, mesmo com um valor significativo para o erro de reconstrução, ou seja recorrendo a um escasso número de componentes.

(Barlow 1989); (Edelman et al. 2001); (Edelman, Intrator, and Jacobson 2002);

7.1 Modelo de memória

Este modelo assenta no pressuposto de que existem duas memórias de imagens, neste caso de retratos de rostos: uma memória de longo prazo constituída por componentes obtidos a partir de um dicionário de imagens e uma memória de curto prazo constituída por componentes construídos a partir de um novo conjunto de retratos. Este é no entanto um modelo completamente especulativo, pois apesar de haver evidências de dois ou mais níveis de memória nos humanos, não temos qualquer evidência de que este modelo corresponda a mecanismos reais da memória humana. Permite-nos no entanto visualizar, desenvolver e testar um modelo e algoritmos, onde é necessário manter e integrar dois grupos de retratos.

Neste modelo, representado na Figura 133, coloca-se o dilema de decidir se se devem incluir novos componentes da memória de curto prazo na memória de longo prazo, ou não para memorizar novos retratos. Esta decisão poderá ser tomada com base numa estimativa de uma função de custos apresentada em (58).

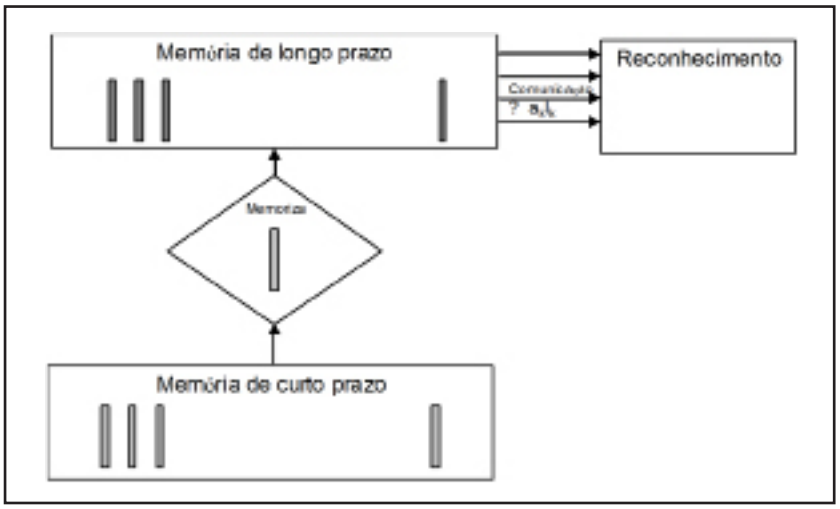


Figura 133 – Modelo de memória de retratos

$$C = C(\text{Memória}) + C(\text{Comunicação}) + C(\text{Erro}) \qquad (58)$$

Esta função de custos vai incluir três componentes: um custo de memória, um custo em comunicação e um custo de erro de construção. Ao decidir incluir novos componentes na memória de longo prazo haverá um custo positivo para essa memória, pois terá de armazenar novos componentes; para a comunicação ($\Sigma a_k I_k$), que terá também de ter em conta um maior número de componentes, o custo poderá no entanto ser positivo ou negativo, obrigando a um estudo aprofundado do valor a entropia no processo de reconstrução das novas imagens; o custo do erro é um custo mais subjectivo, pois um erro grande na reconstrução poderá significar a incapacidade de reconhecer uma face.

7.2 Algoritmo inicial

Para testar o algoritmo inicial utilizámos na memória de longo prazo os componentes PCA que obtivemos a partir da nossa BD com 400 imagens. Na memória de curto prazo usámos um grupo de 39 imagens novas e calculámos os respectivos componentes PCA e ICA. Analisámos depois a evolução do erro no processo de reconstrução das novas imagens, à medida que incluíamos na memória de longo prazo mais um novo componente. Comparámos vários cenários.

Começámos por testar dois critérios para a ordenação dos componentes da memória de curto prazo a mover:

1) No primeiro critério era escolhido sempre o componente na memória de curto prazo que estivesse mais distante de todos os componentes da memória de longo prazo, assumindo que um componente mais distante poderia ter uma contribuição maior para a reconstrução de novos retratos. A distância entre componentes é medida pelo comprimento de cada vector diferença U1-U2 (no espaço de dimensão N a que os vectores componentes pertencem)
2) No segundo critério era escolhido o componente com maior peso na reconstrução dos novos retratos. A importância de cada componente é medida pelo valor médio do valor absoluto dos coeficientes associados a esse componente, na matriz de mistura A, calculada pelo algoritmo *pesqproj*.

O algoritmo *pesqproj*, descrito em 4.3.3, foi adaptado para projectar uma matriz de imagens e para calcular o erro de reconstrução de cada imagem e o erro médio final, pois o conjunto de eixos escolhido pode não ser ortogonal. O algoritmo é iterativo e calcula a matriz de mistura A e o erro para cada conjunto de componentes, passo a passo, até movermos todos os componentes, um a um, para a memória de longo prazo.

$$X1 = A1_i * (U1 + V1)_i \qquad (59)$$

$$X2 = B1_i * (U1 + V1)_i \qquad (60)$$

X1 e X2 são as duas matrizes de retratos, $(U1+V1)_i$ é o grupo de componentes PCA(U) e ICA(V) na iteração **i** e $A1_i$ e $B1_i$ são as respectivas matrizes de mistura. U1+V1 tem mais um componente em cada iteração.

Testámos então, com os dois critérios, 3 cenários distintos, utilizando técnicas estatísticas diferentes na memória de curto prazo:

- ICA
- PCA
- PCA+ICA

Na Figura 134 observa-se que o erro na reconstrução das novas imagens, que residiam na memória de curto prazo, varia de uma forma quase linear, a partir do valor inicial de cerca de 5.1, obtido apenas com os 400 componentes que já existiam na memória de longo prazo, tendendo para zero, à medida que se introduzem novos componentes da memória de curto prazo na memória de longo prazo. Observamos na Figura 134 que a evolução é idêntica com os dois critérios. Também obtivemos resultados idênticos com o cenário PCA+ICA na memória de curto prazo.

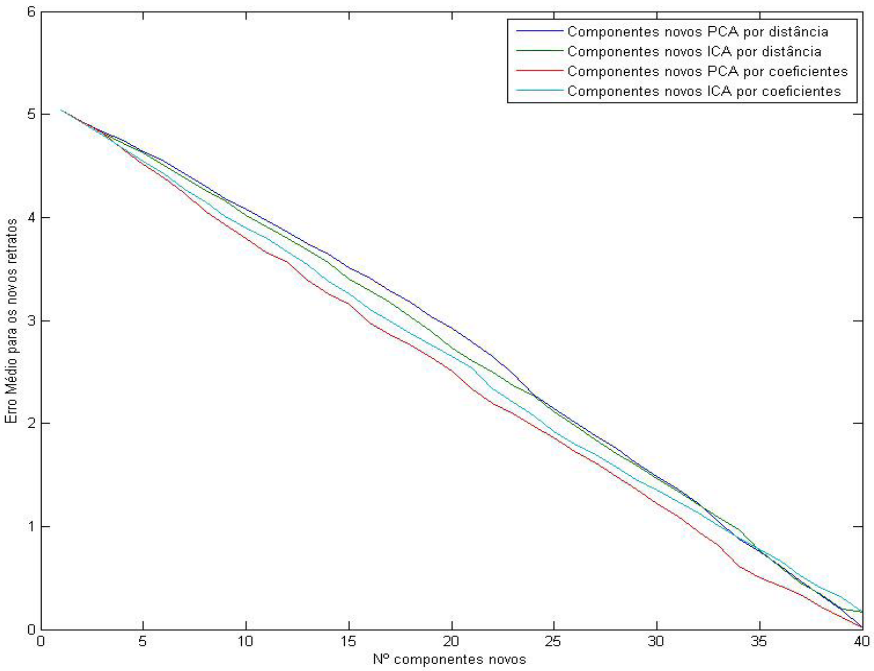


Figura 134 – Variação do erro médio no modelo de memória para novos retratos

A Tabela 21 (critério 1) e a Tabela 22 (critério 2) permitem confirmar que as três curvas, para PCA, ICA e PCA+ICA quase coincidem.

	Erro <5.0	Erro <4.5	Erro <3.0	Erro <1.0
PCA	K= 401	K=406	K=419	K=433
ICA	K= 401	K=406	K=418	K=433
PCA+ICA	K= 401	K=406	K=419	K=434

Tabela 21 – Nº novos componentes (N) para erro < X no modelo dual de memória – critério1

	Erro <5.0	Erro <4.5	Erro <3.0	Erro <1.0
PCA	K= 401	K=405	K=415	K=431
ICA	K= 401	K=405	K=416	K=433
PCA+ICA	K= 401	K=405	K=416	K=434

Tabela 22 – Nº novos componentes (N) para erro < X no modelo dual de memória – critério2

Apesar de a diferença ser muito pequena observa-se que o critério 2 conduziu a uma evolução ligeiramente mais favorável, tendo sido adoptado para o resto do estudo.

Conclusão: observamos que ao mover os componentes para a memória de longo prazo o erro médio global tende para zero, mas a variação do erro é linear e lenta e não muito diferente para todos os cenários; por outro lado este modelo conduz a um crescimento gradual do número de componentes, ou seja mais componentes para cada novo retrato. Justifica-se portanto testar o algoritmo alternativo da secção seguinte.

7.3 Algoritmo com número global de componentes imposto

Como neste caso impomos um número máximo de componentes, igual a 400, para a memória de longo prazo, o algoritmo vai ter de escolher passo a passo os 400 componentes que garantam um melhor resultado para o conjunto dos retratos, antigos e novos. Utilizámos:

- Uma memória de longo prazo:
    - com componentes PCA+ICA – opção mais favorável, de acordo com os resultados das secções 4 e 5, pois permite garantir um erro pré-determinado com menos componentes – sendo os componentes PCA e ICA calculados com os 400 retratos da base de dados
    - com componentes PCA, que reflectem o conhecimento global adquirido dos 400 retratos
- Uma memória de curto prazo:
    - com os componentes PCA
    - com os componentes ICA

As combinações acima permitiram testar 4 cenários distintos. Utilizámos o critério 2, baseado na média dos valor absoluto dos coeficientes, para escolher os componentes a mover e os componentes a substituir. Cada cenário é identificado por uma sigla, por exemplo PCAICA-PCA indica que usamos PCA+ICA na memória de longo prazo e PCA na memória de curto prazo.

7.3.1 Descrição do algoritmo que limita número total de componentes

O algoritmo utilizado generaliza o algoritmo que desenvolvemos para o modelo simplificado

inicial.

a. Começa por calcular as matrizes de mistura A1 e A2, para os dois grupos de imagens

$$X1 = A1 * (U1 + V1) \quad (61)$$

$$X2 = A2 * (U2 + V2) \quad (62)$$

X1 e X2 são as duas matrizes de vectores imagem, calculadas com os dois grupos de retratos. U1+V1 e U2+V2 são respectivamente os componentes PCA(U) e ICA(V), calculados para esses dois grupos de imagens[19]. Para calcular as matrizes A1 e A2 em (61) e (62) utiliza o algoritmo *pesqproj*, adaptado para projectar uma matriz de imagens e para calcular o erro de cada imagem e o erro médio final, pois o conjunto de eixos escolhido pode não ser ortogonal, de acordo com o cenário a testar.

b. Mede a média do valor absoluto dos coeficientes para cada componente na matriz de mistura A1 e escolhe os K componentes U1+V1 com maior peso na reconstrução dos retratos – critério 2. Serão os componentes iniciais a considerar na memória de longo prazo. No nosso caso limitámos o número de componentes a 400 (K=400). Feita esta escolha, o algoritmo calcula A1 e B1 e os erros médios de reconstrução iniciais, com base nos K eixos seleccionados, utilizando uma vez mais uma vez o *pesqproj* generalizado.

$$X1 = A1 * (U1 + V1) \quad (63)$$

$$X2 = B1 * (U1 + V1) \quad (64)$$

c. Teremos assim a matriz A2 e as matrizes A1, B1 iniciais e ainda os dois valores iniciais do erro de reconstrução.

d. Utiliza então um método iterativo para substituir um componente U1+V1 por um componente U2+V2, trocando o componente menos importante em U1+V1 pelo mais importante de U2+V2 – critério 2 – com base na média do valor absoluto dos coeficientes para cada componente em A1 e A2 respectivamente. Para cada iteração **i** o erro e as novas matrizes de mistura A1 e B1 são calculadas, usando uma vez mais o algoritmo *pesqproj*.

$$X1 = A1_i * (U1 + V1)_i \quad (65)$$

$$X2 = B1_i * (U1 + V1)_i \quad (66)$$

Em (65) e (66) $(U1+V1)_i$ é o grupo de componentes seleccionado na iteração *i* e $A1_i$ e $B1_i$ são as respectivas matrizes de mistura.

O processo iterativo termina quando o último componente for substituído.

e. Escolhe o conjunto óptimo de componentes e as respectivas matrizes A1 e B1: o conjunto óptimo de componentes é o que se obtém na iteração em que se garante a menor diferença para os erros médios de reconstrução para os dois grupos de imagens.

19 Cada cenário impõe à partida os U1, V1, U2 e V2 iniciais. Por exemplo no cenário PCAICA-PCA teremos U1 e V1 calculados com o PCA e o ICA para o primeiro grupo de retratos e apenas U2 calculado com o PCA para o segundo grupo de retratos.

f. Calcula a reconstrução de todos os retratos, componente a componente, utilizando o conjunto óptimo de componentes.

$$A = [A1B1] \qquad (67)$$

$$X_{Rk} = A_K * (U1 + V1)^T \qquad (68)$$

A matriz A é a concatenação de A1 e B1. A é a matriz de mistura para o conjunto global de retratos. O vector A_K obtém-se a partir do vector de coeficientes na matriz A que mistura cada retrato X – a linha correspondente em A - anulando os coeficientes de ordem> k, tal como descrevemos em (49). X_{Rk} é o retrato X reconstruído com K componentes. Calcula o erro médio para o conjunto de retratos.

7.3.2 Resultados obtidos

A Figura 135 mostra a evolução do erro de reconstrução médio, calculado com o algoritmo descrito, para os dois grupos de retratos, à medida que se movem componentes da memória de curto prazo para substituir componentes na memória de longo prazo. Observamos que o erro para os novos retratos decresce linearmente de cerca de 5 para zero, tal como tínhamos observado no algoritmo inicial, e que o erro para os retratos antigos aumenta lentamente de um valor inicial entre 0 e 0.7 para um valor final entre 0.2 e 1.1, à medida que substituímos componentes.

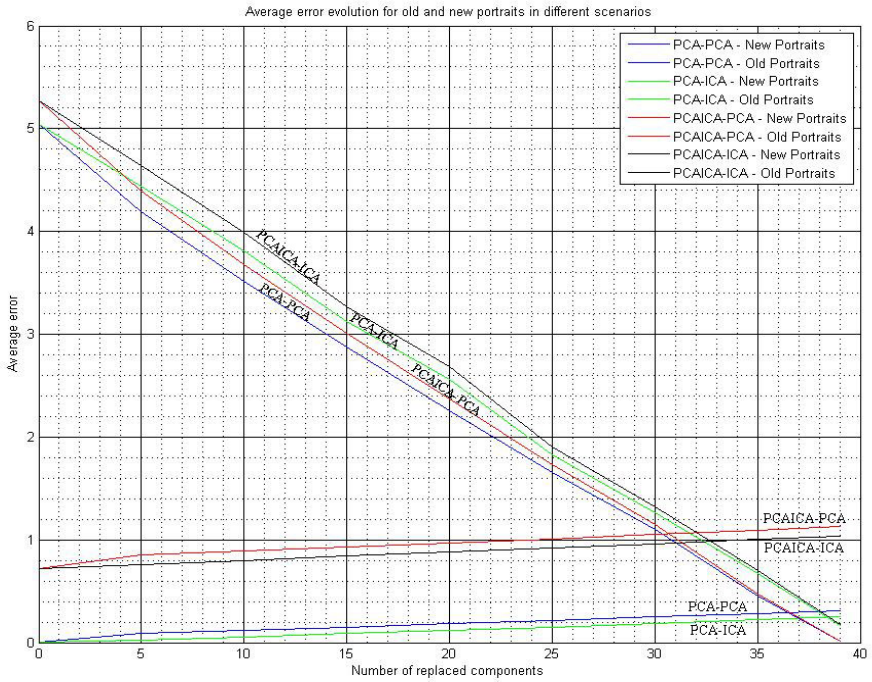


Figura 135 – Variação do erro médio no modelo de memória – 400 componentes

É possível identificar em cada cenário um ponto óptimo, onde as duas curvas de evolução do erro se cruzam, para assegurar o mesmo erro médio para todos os retratos.

Cenário	Nº de componentes a mover	Erro final médio para todos os retratos
PCA-PCA	36	0.30
PCA-ICA	38	0.24
PCAICA-PCA	30	1.05
PCAICA-ICA	32	0.97

Tabela 23 – Condições óptimas em cada cenário

Na Tabela 23 verifica-se que as condições óptimas se observam para 30, 32, 36 e 38 novos componentes, respectivamente para os cenários PCAICA-PCA, PCAICA-ICA, PCA-PCA e PCA-ICA. Com base nesta tabela e na Figura 135 pode concluir-se que o melhor resultado se obtém para o cenário PCA-ICA (curva verde), onde se obteve o erro final médio mais baixo, igual a 0.24, um pouco melhor do que o erro no cenário PCA-PCA. Verifica-se também que o cenário PCAICA-ICA garante um erro um pouco mais baixo do que PCAICA-PCA. Estas diferenças poderão reflectir a complementaridade entre o PCA e o ICA. Mas temos de reconhecer que as diferenças nos 4 cenários são pouco expressivas. Os resultados aparentemente ligeiramente piores, associados aos cenários PCAICA-PCA e PCAICA-ICA, podem ser enganadores, pois resultam apenas de ao inicializar o sistema e escolher os 400 componentes iniciais se impor um erro inicial, ao contrário do que acontece no PCA-PCA e no PCA-ICA. Observamos que os andamentos das curvas são idênticos.

Uma outra forma de comparar os quatro cenários é comparar a evolução do erro de reconstrução para o conjunto global das imagens, tal como se descreveu nos passos e) e f) do algoritmo. Esta reconstrução foi calculada para os 439 retratos com o conjunto óptimo calculado de 400 componentes. A Figura 136 mostra a evolução desse erro para cada um dos cenários. As diferenças são pequenas e utilizámos por isso a escala logarítmica para as evidenciar. PCA-ICA mostra ser uma vez mais a melhor solução.

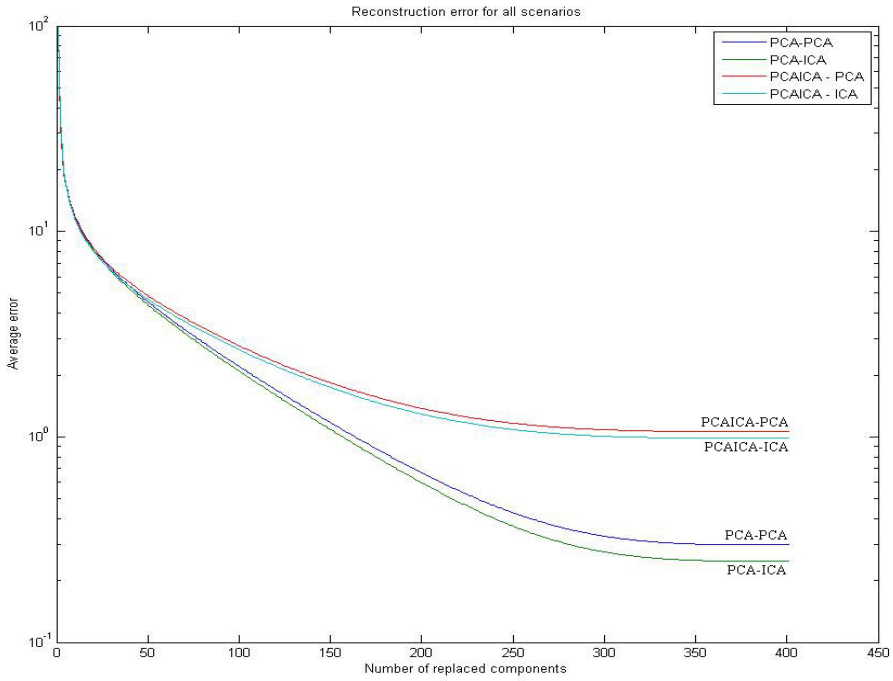


Figura 136 – Variação do erro médio de reconstrução no com o nº de componentes

A Tabela 24 confirma que o cenário PCA-ICA necessita de menos componentes para assegurar um valor pré-fixado para o erro.

Cenário	Erro< 9.0	Erro< 4.5	Erro< 3.0	Erro<1.0	ERRO FINAL
PCA+ICA – PCA	17	55	92	-	1.05
PCA+ICA – ICA	16	52	87	310	0.97
PCA - PCA	18	52	77	163	0.30
PCA-ICA	16	48	74	156	0.24

Tabela 24 – Número de componentes necessários no modelo memória

Será também interessante comparar estes valores com os que se obtiveram para o PCA+ICA e PCA, com todos os componentes, apresentados na secção 5, nas Tabela 12 e Tabela 14, e sintetizados de novo aqui, por facilidade de comparação, na Tabela 25.

Técnica	Erro<9	Erro<4.5	Erro<3	Erro<1	ERRO FINAL
PCA+ICA	K=14	K=37	K=57	K=121	0.0
PCA	K=18	K=52	K=77	K=152	0.0

Tabela 25 – Número de componentes na reconstrução de retratos (ver Tabela 12e Tabela 14)

Pode-se concluir que os valores que obtivemos para este modelo que integra duas BDs de retratos não diferem significativamente dos que obtivéramos, mesmo para o PCA+ICA – a opção mais eficiente que testámos – utilizando todos os componentes PCA e ICA. Com o cenário PCA-ICA obtivemos 16 e 48 componentes em vez de 14 e 37 componentes para assegurar um erro <4.5 e <9 respectivamente. Estes resultados são mesmo melhores que os que tínhamos calculado com o PCA, onde obtivéramos 18 e 52 componentes para as mesmas condições.

Acreditamos que a melhoria, apesar de pequena, que o PCA-ICA demonstra reflecte a complementaridade do PCA e do ICA, possibilitando uma vez mais uma reconstrução mais rápida, necessitando de menos componentes.

A Figura 137 permite comparar a qualidade fotográfica final no cenário PCA-ICA, depois de movido o nº de componentes para a solução identificada como óptima, com os retratos originais.

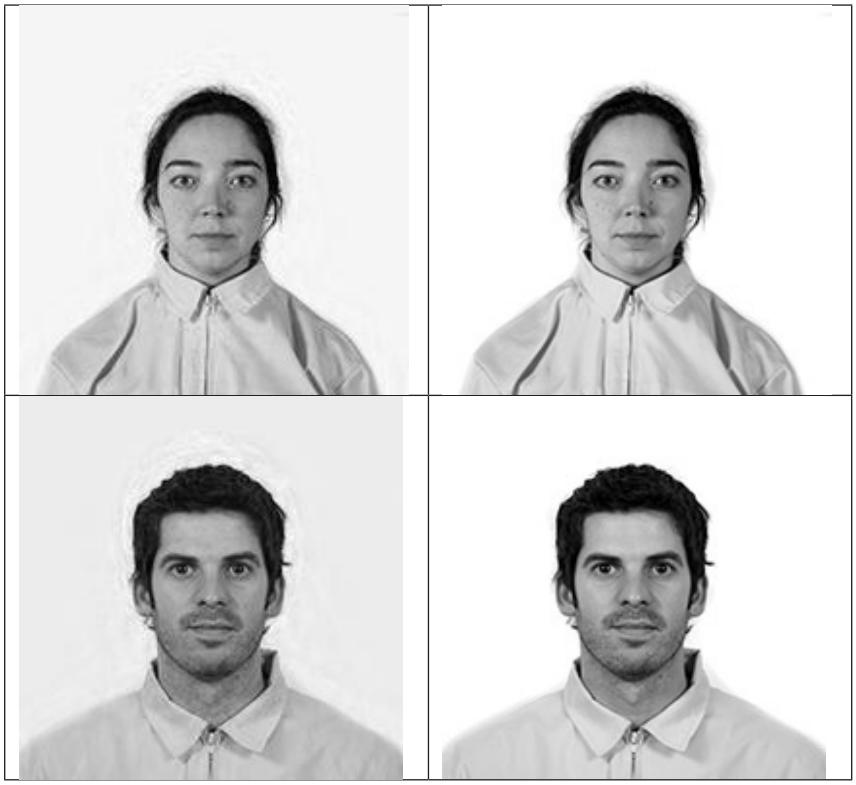


Figura 137 – Retratos reconstruídos com solução óptima

Como se observa o erro médio que calculámos de 0.24 assegura uma qualidade muito boa para os retratos reconstruídos.

7.4 Conclusões

Vamos centrar as conclusões no algoritmo que limita o número total de componentes por ser o mais interessante.

O algoritmo permite reconstruir todos os retratos com valores pequenos para o erro, em todos os cenários testados, mesmo limitando o número total de componentes na memória de longo prazo.

Para o cenário testado com 400 retratos na memória de longo prazo e mais 39 retratos novos, onde se impôs um limite de 400 componentes, conseguiu-se no cenário mais favorável um erro médio global para todos os retratos de 0.24 e em média 16 a 48 componentes para obter um erro de reconstrução que garanta um reconhecimento efectivo a perfeito, de acordo com os resultados do questionário, apresentados na secção 6.

O cenário com melhores resultados é o cenário PCA-ICA, o que acreditamos traduz a complementaridade dos dois métodos. Este cenário permite um menor erro e uma reconstrução com menos componentes.

Os resultados mostram que qualquer cenário testado poderia simular os mecanismos que permitem a aprendizagem não supervisionada de novos retrato, integrando-os no conhecimento de retratos antigos com a recomposição do conjunto de componentes para a reconstrução dos retratos.

Concluímos no entanto que as diferenças entre todos os cenários testados não parecem ser significativas e temos de admitir que esperávamos:

- Uma evolução mais favorável do erro para os novos retratos quando introduzimos novos componentes, ou seja uma curva com evolução mais rápida para zero em vez da evolução linear que observámos. Esperávamos uma curva idêntica às curvas que obtivemos nas secções 4 e 5 para o erro na reconstrução de um retrato. A evolução linear é, acreditamos, o preço pago por termos à partida, só com os componentes da memória de longo prazo, já um erro bastante baixo, cerca de 5.1, e portanto estarmos já na parte da curva perto do valor final, onde de facto a variação é quase linear. Como vimos na secção 6 um erro pouco inferior a 5.1 (4.5 em 100) garante já um reconhecimento quase perfeito.
- Uma diferença maior entre os vários cenários, evidenciando mais a complementaridade do PCA e do ICA

Os testes que realizámos foram limitados pela dimensão da nossa BD – 400 retratos – e da nossa amostra de retratos fora da BD – 39 retratos. Poderia ser interessante testar a evolução dos resultados para os algoritmos desenvolvidos utilizando dimensões diferentes em particular para o número de retratos novos, aumentando-o para o tornar próximo da dimensão da BD original, mas tal implicaria fazer novos retratos ou em alternativa dividir os 439 retratos em dois grupos com a mesma dimensão, por exemplo 220+219, e avaliar a evolução ao fazer a integração de por exemplo grupos de 40 novos retratos até integrar todos os retratos, limitando por exemplo o número de componentes a 220.

O teu corpo é um campo de batalha
Barbara Kruger

OBSERVADOR/OBSERVADO

1. Toda a forma é falsa

2. O tempo actua por debaixo de todas as formas por isso toda a forma é falsa

3. O tempo faz tremer as formas (muda-as)

4. Porém como o tempo passa por baixo de todas as formas – as que são vistas e as que vêem – as formas que tremem vistas por outras formas que tremem tomam aparência, afinal, de formas definitivas, verdadeiras

5. Porém toda a forma é falsa

Gonçalo M. Tavares, em ***Investigações Geométricas***

8 Projecto artístico

O trabalho teórico e de investigação realizado nesta dissertação vai ser utilizado como suporte para projectos fotográficos e artísticos a partir do retrato humano, em torno de questões da identidade. Prevemos realizar vários projectos, tendo como ponto de partida os conceitos desenvolvidos durante este trabalho.

Respondemos já a vários desafios nesta área durante a realização deste trabalho. Foram aceites até à data da conclusão desta disertação propostas de intervenção artística em:

- Ciantec 2009 – III Congresso Internacional de Arte, Novas Tecnologias e Comunicação «Arte, tecnologia e comunicação: novos territórios do conhecimento» de 11/12 a 14 de Outubro de 2009 na Universidade de Aveiro, Departamento de Comunicação e Arte (DCA) – Portugal
- Festival de Arte Feminista para o museu Diogo Sousa em Braga, respondendo a um desafio da UMAR para o Festival de Arte Feminista, o FeministizARTE em Novembro 2009.

Realizou-se, no âmbito desta tese, um primeiro projecto fotográfico, como o nome desta tese, **Espelhos Matriciais**, baseado nos retratos dos alunos e dos professores do Mestrado Multimédia, na sua transformação e reconstrução, utilizando as técnicas estatísticas que exploramos no nosso trabalho. Esse trabalho, mostrado numa sala da FEUP durante algumas semanas posteriores às provas públicas asociadas a esta tese de dissertação, contém um conjunto de fotografias que ilustram o processo de aprendizagem de retratos humanos, como se pode fazer a sua reconstrução com componentes PCA, ICA ou NMF e a forma com nós podemos identificar e reconhecer as identidades das pessoas que nos rodeiam.

Este projecto fotográfico foi orientado pelo Dr. João Fernandes de Serralves e resultou em:

1. Um painel de 100x100 cm com todos os retratos realizados para construir a BD de retratos
2. Um painel de 100x100 cm com todos os componentes PCA
3. Um painel de 100x100 cm com todos os componentes ICA
4. Um painel de 100x100 cm com todos os componentes NMF
5. Um painel de 150x150 cm com imagens reconstruídas com valores diferentes para o erro de todos os elementos ligados ao MM que foram fotografados
6. Um conjunto de 10 painéis 30x90 cm com os retratos de todos elementos ligados ao MM que foram fotografados, em grupos de três, e uma imagem do retrato médio por cada um destes painéis
7. Um vídeo que ilustra o a manipulação dos rostos com estas ferramentas

Para o projecto fotográfico **Espelhos Matriciais** na FEUP foi realizado um desdobrável, com um texto explicativo do processo e com o texto escrito para o projecto pelo Dr. João Fernandes. Foi também impresso um cartaz A3 para divulgar a exposição.

Em anexo, na secção 10.5, apresentamos o texto escrito pelo Dr. João Fernandes para o projecto fotográfico **Espelhos Matriciais** .

Para a realização do vídeo foi escrito um script Matlab que constrói artificialmente retratos para fazer o morphing entre 2 retratos distintos.

Tal como vimos a matriz de retratos pode ser reconstruída com base num conjunto de coeficientes A_{ij} de uma matriz de mistura A que o PCA ou o ICA calculam. No nosso caso cada retrato tem 400 coeficientes PCA, por exemplo, que podem ser vistos como um vector unitário num espaço 400-dimensional. Se pensarmos em 2 retratos distintos vamos ter dois vectores: um vector A1 e um

vector A2, como se observa na Figura 138.

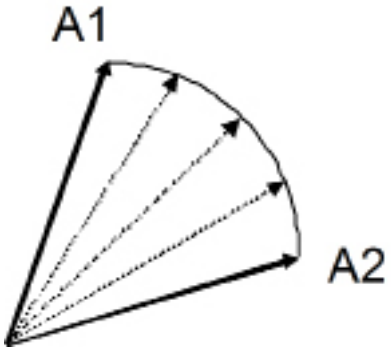


Figura 138 – Cálculo de vectores de mistura para morphing

Para construir os retratos para morphing, o que o algoritmo faz é calcular vectores intermédios entre A1 e A2 e reconstruir retratos virtuais intermédios, entre X1 e X2, utilizando os vectores intermédios de coeficientes na equação (49), fazendo a reconstrução com todos os componentes: K=400. O algoritmo permite criar o número de retratos intermédios pretendido.

Este projecto fotográfico pôde ser visitado na FEUP durante um período de tempo que cobriu a apresentação desta dissertação.

(Roque 2009);(Roque 2009);

A maioria, talvez todos, dos nossos espelhos são imprecisos e rudes, ainda que em grau diverso e de maneiras diferentes. Alguns ampliam, alguns reduzem, outros devolvem imagens lúgubres, cómicas, ridículas, terríveis.

W.H. Auden

Cada rosto parece
o que cada coração
padece

René Ronce [tradução de Renato Roque]

9 Conclusões

Esta secção faz uma listagem sucinta das principais conclusões do trabalho. As conclusões associadas ao NMF e aos algoritmos que apresentámos na secção 7, pela sua especificidade, estão agrupadas em secções independentes.

Tal como acontece sempre ao longo de todo este documento, nestas conclusões, o erro na reconstrução de um retrato é calculado com a equação (48) e é expresso numa escala de 0 a 100.

Temos também consciência de que neste trabalho ficam naturalmente em aberto inúmeras questões, algumas das quais procurámos identificar ao longo do documento. Gostaríamos em particular de ter podido desenvolver e testar conceitos como o de quantidade de informação de um novo retrato e de entropia nos retratos humanos e de ter mapeado medidas desse conceito para um retrato com os algoritmos que apresentámos na secção 7, para integrar duas BDs de retratos, tendo como inspiração um modelo dual de memória, associado a um modelo de aprendizagem não supervisionada de novos retratos.

9.1 Conclusões gerais

Os testes por nós realizados com a nossa BD de retratos com o PCA, ICA e PCA+ICA permitem-nos concluir que:

a. Todas as técnicas permitem uma reconstrução perfeita, sem erro, para além do erro de cálculo – obtivemos erros da ordem de 1 e-025 – de qualquer retrato da BD de retratos, com os componentes calculados a partir dessa BD de retratos
b. A curva do erro de reconstrução converge rapidamente para zero e poucos componentes são necessários para garantir valores para o erro que possibilitam o reconhecimento do retrato
c. A reconstrução que utiliza menos componentes é a que é baseada em PCA+ICA, que parece tirar partido da complementaridade do PCA e do ICA, para permitir uma reconstrução mais rápida
d. Todas estas técnicas permitem também reconstruir retratos novos com um valor para o erro que depende da dimensão da BD e da resolução dos retrato – com 400 retratos 200x200 pixéis o valor para o erro é de 5.1 e permite um reconhecimento razoável dos retratos
e. A curva do erro de reconstrução para retratos novos também converge rapidamente e poucos componentes são necessários para ter valores próximos do erro final;
f. Obtiveram-se curvas que mostram a relação entre a resolução dos retratos e a dimensão necessária para a BD, para garantir um determinado valor para o erro na reconstrução de retratos novos
g. O PCA+ICA é também a técnica que possibilita uma reconstrução mais rápida dum retrato novo.

Recorrendo à normalização da BD, fazendo coincidir os dois olhos de todos retratos, verificámos que:

a. As curvas de reconstrução dos retratos da BD, apenas convergem um pouco mais rapidamente para o PCA e para o PCA+ICA;
b. No caso do ICA as duas curvas coincidem;
c. Para o PCA e PCA+ICA consegue-se garantir o mesmo erro com um pouco menos componentes, mas a diferença é relativamente pequena;
d. Verifica-se que a normalização tem um efeito no PCA mas parece não ter qualquer efeito

no ICA;

e. Para retratos novos o efeito da normalização nos casos do PCA e PCA+ICA é já significativo. O erro final diminui mas sobretudo conseguem-se os mesmos valores para os erros com menos de 50% dos componentes.

O questionário que apresentámos na secção 6 permite-nos tirar um conjunto de conclusões que nos parecem importantes:

a. Quando o erro geométrico tem valores superiores a 10.5 o reconhecimento parece ser muito difícil;
b. Quando o erro desce para valores próximos de 9 o reconhecimento torna-se efectivo, com taxas de sucesso da ordem dos 50%;
c. Quando o erro é inferior a 4.5 o reconhecimento parece ser quase perfeito.

Tendo em conta as conclusões anteriores – em particular os dois limiares do erro – e os resultados que obtivemos nas secções 4 e 5 podemos afirmar que utilizando uma BD com 400 retratos com retratos de 200x200 pixéis conseguimos:

a. Um reconhecimento entre efectivo e perfeito de um retrato da BD com 18 a 45 componentes com o PCA+ICA;
b. Um reconhecimento efectivo de um retrato novo com 48 componentes com o PCA+ICA.

Normalizando a BD teremos valores mais favoráveis:

a. Um reconhecimento dos retratos da BD entre efectivo e perfeito de um retrato da BD com 14 a 37 componentes com o PCA+ICA, em vez dos 18 a 45 anteriores;
b. Um reconhecimento efectivo de um retrato novo com cerca de 20 componentes com o PCA+ICA, em vez dos 48 anteriores.

Pode-se calcular a evolução dos números anteriores apresentados, quando se altera a resolução das imagens e a dimensão da BD.

Pudemos também obter um conjunto de resultados para o que designámos como erro ou qualidade fotográfica:

a. Verificámos que parece existir uma relação entre o erro geométrico, medido com o módulo do vector erro na reconstrução, e uma qualidade fotográfica subjectiva, observada para diversos valores do erro. Pudemos confirmar que um erro de δ parece sempre corresponder à mesma qualidade fotográfica do retrato reconstruído;
b. Esta relação foi confirmada para as diferentes técnicas estatísticas testadas e comparadas, inclusive o NMF e o PCA+ICA+NMF, para várias resoluções dos retratos, para dimensões diferentes da BD e para retratos que pertenciam à BD ou que estavam fora da BD.

O conjunto das conclusões anteriores permite-nos afirmar que parece ser possível com a BD que utilizámos de 400 retratos e com 200x200 pixéis reconstruir qualquer retrato de face frontal, de qualquer pessoa no mundo, com um erro que permitiria um reconhecimento efectivo (erro=5.1 < 9). As curvas que obtivemos para o erro em função da dimensão da BD e da resolução dos retratos podem mesmo nos levar a estimar que com um número de retratos entre 600 a 800 se conseguiria um reconhecimento perfeito, ou seja um erro < 4.5.

Este resultado parece ser relevante no que diz respeito aos conceitos de **aprendizagem não supervisionada** (*unsupervised learning*) **e codificação esparsa** (*sparse coding*). Relevante para a codificação esparsa porque mostra que poucos componentes possibilitam a reconstrução de qualquer retrato com erros aceitáveis. Relevante para a aprendizagem não supervisionada porque se observa que podemos reconstruir novos retratos com componentes antigas, com erros aceitáveis do ponto de vista do reconhecimento.

9.2 Conclusões para o NMF

Quando iniciámos este trabalho o NMF não estava no nosso horizonte. No entanto sendo uma técnica não supervisionada tal como o PCA e o ICA, parecia ser interessante testá-la e compará-la e sobretudo verificar se o sistema híbrido PCA+ICA+NMF não poderia acrescentar ainda qualquer melhoria ao PCA+ICA. Ao termos encontrado SW que implementava vários algoritmos propostos para o NMF, não hesitámos.

Depois de ler vários artigos e teses que propunham o NMF na área do reconhecimento temos de começar por reconhecer que não nos pareciam de todo bem justificadas as vantagens do método. Não parecia claro a vantagem da positividade dos coeficientes, pois os valores negativos no PCA e ICA resultam sobretudo dos métodos eliminarem as médias e por se focarem na variação ou seja na informação efectiva.

Depois de testar o NMF, recorrendo a 3 algoritmos diferentes que são baseados em métodos diferentes concluímos que:

a. O NMF não parece ser adequado à reconstrução e reconhecimento de retratos;
b. Não obtivemos qualquer evidência de que há uma relação dos componentes NMF com partes específicas do retrato;
c. O erro de reconstrução, mesmo nos melhores casos, é grande e não parece claro que diminua para zero se aumentarmos as iterações, a dimensão da BD ou o número de componentes. Observámos em certos casos que o erro diminui, mas tão lentamente que não tivemos qualquer evidência da convergência para zero;
d. Observámos que existe uma dependência grande das inicializações das matrizes no caso do método GP, o que dificulta a utilização do método. Em todos os casos há um conjunto de parâmetros como a dimensão da BD ou o número de componentes em que não é fácil definir a solução óptima;
e. Verificámos que não conseguimos qualquer melhoria com o sistema PCA+ICA+NMF. Este resultado pode dever-se em parte às contradições entre os métodos no que diz respeito à negatividade de coeficientes e das consequências que resultam na manipulação das médias, como descrevemos na secção 4.6;
f. No caso do GP observámos que a normalização da BD dificultou o NMF, fazendo com que houvesse mesmo retratos onde a reconstrução se tornou impossível. Não tivemos oportunidade de testar outros métodos NMF, em virtude de não termos encontrado relevância no NMF para o foco do nosso trabalho.

9.3 Conclusões sobre algoritmos de integração de duas BDs

Esta secção sintetiza conclusões da secção 7 em que se propuseram e testaram modelos e algoritmos para integrar de uma forma optimizada duas BDs, podendo utilizar técnicas estatisticas diferentes em cada uma das BDs:

a. O algoritmo apresentado permite reconstruir todos os retratos de duas BDs diferentes a integrar com valores pequenos para o erro, em todos os cenários testados, mesmo limitando o número total de componentes para integração;
b. O algoritmo permite optimizar a integração para o conjunto global de imagens;
c. O cenário com melhores resultados é o cenário PCA-ICA, o que acreditamos traduz uma vez mais a complementaridade dos dois métodos. Este cenário permite um menor erro e uma reconstrução com menos componentes;
d. Os resultados mostram que qualquer cenário testado poderia simular os mecanismos que permitem a aprendizagem não supervisionada de novos retrato, integrando-os no

conhecimento de retratos antigos, com a recomposição do conjunto de componentes para a reconstrução dos retratos

(Roque, Puga, Carrapatoso et al. 2009); (Roque, Puga, Barbosa et al. 2009);

Fotografia é o nosso exorcismo. Sociedades primitivas tinham as máscaras, a sociedade burguesa o espelho. Nós temos as nossas imagens.

Jean Beaudrillard

Há varandas em todo olhar e os que habitam vêm passear-se nelas, vêm ver quem passa, ou respirar, ou simplesmente descansar da tarefa de estar dentro, fumar os ares da cidade, sacudir os lençóis de casal, reclamar a pertença das fachadas. Poucos vão à varanda. Poucos vêm à varanda. E menos são os que vêm à varanda como quem decididamente vem à varanda. E menos ainda são os que quando vêm à varanda saem realmente, vêm de facto cá fora. Por isto, muitas varandas são enganos, são só promessas de exterior, não são mais que pequenas extensões visíveis do dentro.
Poucos, quase nenhuns, vêm à varanda com a devida importância. É importante fazer as coisas com a devida importância. Uma varanda exterior permite esse milagre do solo firme para o ver de voo. Mas muitos, quase todos, quando vêm à varanda não vêm à varanda. Estão ainda no sofá, no banho, nos afazeres da escrivaninha. Ou então nem em casa estão. Saíram há muito tempo e a verdade é que não voltaram. Nunca regressaram. Ainda assim podemos vê-los à varanda. Eles é que não podem ver-nos. Mas há outros. Há aqueles que sabem bem dos privilégios de se ter uma varanda, têm muito presente o acrescido valor que, nos dia que correm, com a agoniante falta de lugar, representa poder estar num espaço exterior, uma real alternativa ao cansaço das paredes, das cadeiras, ao peso dos tectos, ao desarticulado arfar das alcatifas.
Umas são de ferro, umas de cimento, outras estão talhadas na própria pedra dos prédios. Essas parecem predicar o privilégio. Algumas são austeras, outras simplesmente geométricas. Muitas reservam, além da ameaça do desmoronamento, e não sei bem porquê, um ar clássico. Erguem-se em colunas, desenham-se em capitéis e terminam nos cotovelos de olhar as ruas, quase cornucópias sem natureza.
Imagino um mundo sem varandas. Toma-me uma tristeza absoluta. Compadeço-me, contra a minha vontade, dessa visão. Comovo-me. Olho outra vez todas as fachadas ondulantes, toda a arquitectura se desfaz num bocejo. Perco momentaneamente o ar. Venho à varanda. Respiro. Volto as costas ao interior. Posso agora, graças ao maravilhoso invento da perspectiva, ver todas as varandas desenhadas sobre as linhas de fuga. Quase todas estão vazias. Que pena!
Devem estar todos muito ocupados a habitar o interior.

Marta Bernardes

10 Anexos

10.1 Endereços URL

Listagem de endereços URL com as BDs que avaliámos para utilizar no nosso trabalho, acabando por decidir utilizar a faces94, enquanto não dispúnhamos de BD do projecto

http://www.face-rec.org/databases/ (Junho 2009)
http://face.nist.gov/colorferet/ (Junho 2009)
http://cvc.yale.edu/projects/yalefaces/yalefaces.html (Junho 2009)
http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html (Junho 2009)
http://cbcl.mit.edu/software-datasets/heisele/facerecognition-database.html (Junho 2009)
http://cswww.essex.ac.uk/mv/allfaces/index.html (Junho 2009) (Junho 2009)
http://cswww.essex.ac.uk/mv/allfaces/faces94.html (Junho 2009) (Junho 2009)
http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html (Junho 2009)

Listagem de endereços URL com as conferências referidas no trabalho

AFGR - the IEEE conference series on Automatic Face and Gesture Recognition
http://www.fg2008.nl/ (Junho 2009)

AVBPA – Audio and Video based Biometric Person Authentication conference
http://www.iapr.org/members/newsletter/Newsletter06-01/index_files/Page571.htm (Junho 2009)

Listagem de endereços URL com o SW utilizado no trabalho

FastICA

http://www.cis.hut.fi/projects/ica/fastica/code/dlcode.shtml (Junho 2009)

Toolbox da Universidade Técnica da Dinamarca

http://isp.imm.dtu.dk/toolbox/index.htm (Junho 2009)

Algoritmo dos gradientes projectados de Chih-Jen Lin

http://www.csie.ntu.edu.tw/~cjlin/nmf/ (Junho 2009)

RScorePlus SW - Parameter Estimation of Signal Detection Models

http://psych.colorado.edu/~lharvey/html/software.html (Junho 2009)

10.2 Referências

Abdi, Hervé. 1987. *Introduction au traitement des donnés expérimentales*: Presses Universitaires de Grenoble.

———. 2007. Encyclopedia of Measurements and Statistics.

Adolphs, Ralph. 2003. Cognitive neuroscience of human social behaviour. *Nature reviews* 4:165-178.

Almeida, Helena. 1995. *Dramatis Persona: Variações e fuga sobre um corpo* Fundação de Serralves.

Alves da Silva, Augusto, Daniel Blaufuks, Susanne Themlitz, and Vasco Araújo. 2006. Catálogo Bes Photo 2006 CCB.

Attick, J., and A. Redlich. 1992. What does the retina know about natural scenes. *Neural Computation* 4:196-210.

Attneave, F. 1954. Some informational aspects of visual perception. *Psychological Review*, 183-193.

Barbeau, Emanuel, Margot Taylor, Jean Regis, Patrick Marquis, Patrick Chauvel, and Catherine Chauvel. 2008. Spatio temporal Dynamics of Face Recognition. *Cerebral Cortex* (May 2008:18):997-1009.

Barlow, Horace. 1989. Unsupervised learning. *Neural Computation* (1):295–311.

———. 2001. The Exploitation of Regularities in the Environment by the Brain. *Behavioural and Brain Sciences* 24

Barthes, Rolland. 1980. *La chambre claire*: Gallimard.

Bartlett, Marian Stewart. 1998. Face image analysis by unsupervised learning and redundancy reduction. Ph.D., University of California, San Diego.

Berry, Michael W., Murray Browne, Amy N. Langville, V. Paul Pauca, and Robert J. Plemmons. 2007. Algorithms and applications for approximate nonnegative matrix factorization. *Computational Statistics & Data Analysis* 52 (1):155.

Blaufuks, Daniel. 2007. *Sob céus estranhos*: Tinta da China.

Bolme, David, J. Ross Beveridge, Marcio Teixeira, and Draperm Bruce A. 2003. The CSU Face Identification Evaluation System: Its Purpose, Features and Structure. In *International Conference on Vision Systems,*. Graz, Austria.

Boris Lehman. 2003. *Histoire de ma vie racontée par mes photographies*: Centre George Pompidou.

Bruce, Vicky, and Andy Young. 1986. Understanding Face Recognition. *British Journal of Psychology* (77):305,327.

Chih-Jen, Lin. 2007. Projected Gradient Methods for Non-negative Matrix Factorization. *Neural Computation* 19:2756-2779.

Draper, B., K. Baek, M.S. Bartlett, and R. Beveridge. 2003. Recognizing Faces with PCA and ICA. *Computer Vision and Image Understanding* Volume 91 (Issues 1-2):115-137

Edelman, S., B. Hiles, P.H. J. Yang, and N. Intrator. 2001. Probabilistic principles in unsupervised learning of visual structure: human data and a model. Paper read at 15th Annual Conference on Neural Information Processing Systems (NIPS), Dec 03-08, at Vancouver, Canada.

Edelman, S., N. Intrator, and J. S. Jacobson. 2002. Unsupervised learning of visual structure. Paper read at 2nd International Workshop on Biologically Motivated Computer Vision (BMCV 2002), Nov 22-24, at Tubingen, Germany.

Ekman, Paul. 1999. *Handbook of Cognition and Emotions*. John Wiley and Sons Ltd ed. Vol. Facial Expressions.

Fontcuberta, Joan. 1988. *Fauna Secreta*.

Friedlander, Lee. 1970. *Self Portrait*.

Frizot, Michel. 1998. *A new history of Photography*: Konemann.

Gross, C. G., and J. Sergent. 1992. Face recognition. *Current Opinion in Neurobiology* 2: 156-161.

Gross, Charles G. 2002. Genealogy of Grandmother Cell. *The Neuroscientist*.

Gross, Ralph, Jiambo Shi, and Jeff Cohn. 2001. Quo Vadis Face Recognition? *Robotics Institute*.

Harvey, Lewis O. 2009. RscorePlus user's manual - Parameter Estimation of Signal Detection Models: Department of Psychology and Neuroscience - University of Colorado.

Heeger, David. 2003. Signal Detection Theory.

Hyvärinen, Aapo. 1999. Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. . *IEEE Transactions on Neural Networks* 10 (3):626-634.

———. 1999. Survey on Independent Component Analysis. *Neural Computing Surveys* Vol. 2: pp. 94-128.

Hyvärinen, Aapo , Erkki Oja, and Karhunen Juha. 2001. *Independent Component Analysis*: Wiley.

Hyvärinen, Aapo, and Erkki Oja. 2000. Independent Component Analysis Algorithms and Applications. *Neural Networks* 13 ((4-5)):411-430.

Jackson, J. E. 1991. *A User's Guide to Principal Components*: John Wiley and Sons, .

Jolliffe, I.T. . 2002. *Principal Component Analysis* Ed Springer

Lee, Daniel, and H. Sebastian Seung. 2001. Algorithms for non-negative matrix factorization. *Adv. Neural Info in Processing Systems* 13 556-562.

Maddow, Ben. 1977. *Faces. A narrative history of the portrait in photography*: New York Graphic Society, Boston.
Michaels, Duane. 1997. *The essential Duane Michaels*: Bullfinch.
Molder, Jorge. 1999. *Luxury Bound*: Assírio & Alvim.
Neto, José Luís. 2003. *22474*: Assírio e Alvim.

Puga, André. 2000. Análise e compressão de imagem usando componentes independentes, simétricas e principais, Porto.
———. 2001. A computational allegory for V1. Paper read at 2nd International Symposium on Image and Signal Processing and Analysis at PULA, CROATIA.
Roque, Renato. 1986. Entre 2 espelhos a 10º. Porto.
———. 1988. 10º para a meia-noite. Porto.
———. 1992. *Espelhos*.
———. 1992. Novos Espelhos. Porto.
———. 2009. Matrix Mirrors. Porto: UT Austin|Portugal Program – Submitted to Future Places.
———. 2009. Mattrix Mirrors. Aveiro: Submitted to III Congresso Internacional de Arte, Novas Tecnologias e Comunicação.
Roque, Renato, André Puga, Fernando Barbosa, and Pedro Almeida. 2009. How well can we recognize facial portraits reconstructed with a few statistical components? Submitted to *Spatial Vision*.
Roque, Renato, André Puga, Eurico Carrapatoso, and Fernando Barbosa. 2009. On building photographic portraits using statistical techniques. Paper at VIP 2009, at Porto.
Sander, August, and John Von Hartz. 1991. *August Sander (Aperture Masters of Photography)*: Aperture.
Shannon, Claude. 1948. Mathematical Theory of Communication. *The Bell System Technical Journal* Vol. 27, :pp. 379–423, 623–656.
Sherman, Cindy, and Peter Galassi. 2003. *The Complete Untitled Film Stills* The Museum of Modern Art, New York
Tovée, Martin. 1995. What are faces for? *Current Biology* 5 (5):480-483.
Trehub, Arnold. 1997. Sparse Coding of Faces in a Neuronal Model: Interpreting Cell Population Response in Object In *Neural Network Models of Cognition*.
Valadas, Cristina. 2008. Mente Semente: Gémeo R.
Vieitez, Virxilio. 1998. *Album*: Fundacio Caixa Galicia.
Wendy S. Yambor, Bruce A. Draper and J. Ross Beveridge. 2002. Analyzing PCA-based Face Recognition Algorithms: Eigenvector Selection and Distance Measures. *World Scientific Press*
Wickens, Thomas D. 2001. *Elementary Signal Detection Theory*: Oxford University Presss.
William, Ewing. 2006. *Face. The new photographic portrait*: Thames & Hudson.
Xudong, Xie. 2006. Face image analysis and its applications. Ph.D., Hong Kong Polytechnic University (Hong Kong), Hong Kong.
Yambor, Wendy S., Bruce A. Draper, and J. Ross Beveridge. 2000. Analyzing PCA-based Face Recognition Algorithms: Eigenvector Selection and Distance Measures. *World Scientific Press*
Yang, Ming-Hsuan. 2002. Detecting Faces in Images: a Survey. *IEE Transactions on Pattern Analysis and Machine Intelligence* 24 (1).
Young M. P., Yamane S. 1992. Sparse population coding of faces in the inferotemporal cortex. *Science* (256):1327-1331.
Yun, Xue. 2007. Non-negative matrix factorization for face recognition. Ph.D., Hong Kong Baptist University (Hong Kong), Hong Kong.
Zafeiriou, S., A. Tefas, I. Buciu, and I. Pitas. 2005. Class-specific discriminant non-negative matrix factorization for frontal face verification. Paper read at 3rd International Conference on Advances in Pattern Recognition, Aug 22-25, at Bath, ENGLAND.
Zafeiriou, S., A. Tefas, and I. Pitas. 2005. Discriminant NMFFaces for frontal face verification. Paper read at IEEE Workshop on Machine Learning for Signal Processing (MLSP), Sep 28-30, at Mystic, CT.
Zhao, W., R. Chellappa, P. J. Phillips, and A. Rosenfeld. 2003. Face recognition: A literature survey. *ACM Computing Surveys* 35 (4):399-458.

10.3 Conceitos da teoria da informação

Se X é uma variável aleatória que pode ter um conjunto de valores finitos, de acordo com uma distribuição de probabilidades p(X), define-se quantidade de informação associada a uma ocorrência x_i da variável como

$$I(x_i) = \log_2 1/p(x_i) \qquad (1)$$

A quantidade de informação mede aquilo que passamos a saber, que não sabíamos, quando tomamos conhecimento dessa ocorrência. Mede-se em bits. Pela expressão constata-se que, como seria de esperar, a informação aumenta quando a probabilidade diminui. A informação associada a um evento muito pouco provável é bastante maior do que a associada a um evento com grande probabilidade de ocorrer. Por exemplo se a probabilidade for ½ a informação I será 1 bit, se a probabilidade for ¼ a informação será de 2 bits.

Um outro conceito fundamental é o conceito de entropia. A entropia de uma variável aleatória X mede da quantidade de informação necessária para, em média, descrever a variável

$$H(X) = \Sigma\ p(x_i) \log_2 1/p(x_i) \qquad (2)$$

$$0 \leq H(X) \leq \log_2 n \qquad (3)$$

A expressão (3) representa o valor máximo da entropia, sendo n o nº de acontecimentos possíveis, ou seja, por exemplo, para 256 acontecimentos possíveis, a entropia é no máximo 8, o que significa que precisamos de no máximo 8 bits para os codificar.

Define-se ainda entropia conjunta das variáveis. X e Y como

$$H(X,Y) = \Sigma\ p(x_i , y_j) \log_2 1/ p(x_i , y_j) \qquad (4)$$

Se X e Y são independentes – o que significa que os valores de uma não são influenciados pelos valores da outra - então:

$$H(X,Y) = H(X) + H(Y) \qquad (5)$$

Entropia conjunta é então a medida da quantidade de informação necessária para, em média, descrever um par de variáveis aleatórias

Um outro conceito da teoria de informação, fundamental para Barlow, na teoria que desenvolveu a partir do trabalho de Attneave, para explicar o processo de aprendizagem automática, como veremos na secção seguinte, é o conceito de codificação mínima.

Suponhamos que temos um texto escrito em português. Utilizando a codificação ASCII, teremos necessidade de 7 bits por carácter – o que nos permite ter 128 caracteres diferentes – e poderíamos esperar, por isso, que a entropia de um carácter seria de 7 bits, mas tal não acontece.

De facto, tomando C como a variável (carácter) que pode assumir um de 128 valores possíveis (7 bits), teremos, recorrendo à expressão (2)

$$H(C) = \Sigma\ p(c_i) \log_2 1/p(c_i) \qquad (6)$$

Se a probabilidade de ocorrência de cada carácter fosse 1/128 e se a probabilidade de ocorrer um carácter não dependesse da ocorrência de outros caracteres, calculando H obteríamos o valor esperado 7. De facto, nesse caso, as 128 parcelas valeriam cada uma

$$1/128 \log_2(128) = 1/128*(7) = 7/128 \quad (7)$$

e portanto o somatório das 128 parcelas seria igual a 7. O valor de H é na realidade menor do que 7, em virtude de termos caracteres muito prováveis e outros muito pouco prováveis. Observou-se que, dependendo do texto analisado, se encontra um valor de entropia de 1ª ordem, para um carácter de texto, entre 4 e 5 bits.

Se analisarmos este problema pela perspectiva dos bits, verificamos que a entropia de cada um dos bits, b_i, vai ser, recorrendo à mesma expressão (6), pois o bit só pode ter dois valores: 0 e 1, com probabilidades $p(b_i)$ e $(1-p(b_i))$

$$H(b_i) = p(b_i) \log_2 1/p(b_i) + (1-p(b_i)) \log_2 1/(1-p(b_i)) \quad (8)$$

Se utilizamos um código de 7 bits, para representar cada carácter, somando o valor da entropia para todos os bits, obtemos mais uma vez, necessariamente, um valor menor que 7, que é o valor máximo. E é menor do que 7, em virtude de mais uma vez a probabilidade dos valores 0 e 1 dos diferentes bits não ser a mesma e dos vários bits poderem estar correlacionados entre si – se todas as probabilidades fossem ½, calcula-se facilmente o valor da entropia dos 7 bits, e obtemos o valor 7.

De acordo com a teoria de Shannon é possível encontrar sempre códigos de comprimento variável – em vez de um código de comprimento fixo, com o valor da entropia máxima – com comprimentos médios iguais à entropia, ou quando muito ao valor da entropia +1

$$H \leq L < H + 1 \quad (9)$$

implementando assim um código que minimiza o somatório das entropias dos bits e portanto corresponde à entropia real do conteúdo da informação. No exemplo utilizado, do texto escrito em português, seria portanto possível encontrar um código de comprimento variável, com em média 5 ou 6 bits, em vez dos 7 bits do código ASCII.

Este conceito genérico de entropia e de código optimizado pode ser generalizado. Desta forma estamos a reduzir a redundância e a aumentar a independência na informação. De acordo com as teorias de Barlow um mecanismo idêntico, que permite descodificar imagens num conjunto de componentes independentes, é essencial para possibilitar a aprendizagem, em particular para identificar relações e para identificar o novo naquilo que seria expectável, mecanismos fundamentais para aprender de uma forma automática a partir da percepção de imagens.

(Shannon 1948);

10.4 Inquérito

Nesta secção apresenta-se o inquérito que enviámos a todas as pessoas – alunos e professores do MM – e de que recebemos 28 respostas, que utilizámos na análise que apresentamos na secção 6.

Apresentamos também o texto que enviámos a apresentar o inquérito.

Este inquérito insere-se num trabalho de investigação, no âmbito da minha tese de mestrado, que tem como um dos objectivos determinar o grau de reconstrução necessário para a identificação de pessoas através de fotografias, construídas com ferramentas estatísticas aplicadas a imagens de uma Base de Dados.

Perante cada retrato, identificados de F1 a F48, quero que digam **se o reconhecem ou não** e, se o reconhecerem, que digam **quem é essa pessoa** e **qual o grau de certeza** com que fazem o reconhecimento: um grau de certeza pequeno (**P**), razoável (**R**) ou absoluto (**A**).

O inquérito não vos vai tomar mais do que alguns minutos e é muito importante para eu poder analisar a qualidade subjectiva dos retratos reconstruídos. Acho que poderá ter inclusive um carácter lúdico e de jogo divertido, ao tentar identificar conhecidos e amigos nas imagens algo fantasmagóricas que preenchem as folhas do inquérito.

Tenham em conta que na lista há fotos **de pessoas do Mestrado Multimédia**, que conhecem bem, fotos com **diferentes níveis de qualidade**, mas **também há retratos de pessoas que desconhecem**. Um dos objectivos é identificar uns e outros.

Todos os dados individuais serão tratados de forma absolutamente anónima e confidencial; apenas os dados estatísticos de grupo serão publicados na tese e ficarão disponíveis para quem os quiser conhecer.

Apresento um pequeno exemplo para não haver dúvidas sobre o preenchimento, mas se houver, não hesitem em contactar-me.

Mesmo aqueles que não foram fotografados por mim podem responder a este inquérito.

Agradeço a vossa colaboração, que é essencial para poder obter dados estatísticos fundamentais para o meu trabalho.

Obrigado

Renato Roque

EXEMPLO de PREENCHIMENTO

FOTO A	FOTO B

	Foto A	Foto B
Reconhece a pessoa? (S/N)	S	N
Se reconhece, quem é? (escreva o nome da pessoa)	Cavaco	-
Grau de certeza no reconhecimento (Absoluta (A)/ Razoável (R)/ Pequena (P))	A	-

Nota: neste exemplo com duas fotos, numa reconheço o Prof. Cavaco Silva com certeza **A**bsoluta, na outra não reconheço o personagem apresentado.

As páginas seguintes listam os retratos para os quais queria as vossas respostas.

FOTO1	FOTO2	FOTO3	FOTO4
FOTO5	FOTO6	FOTO7	FOTO8

	Foto 1	Foto 2	Foto 3	Foto 4	Foto 5	Foto 6	Foto 7	Foto 8
Reconhece a pessoa? (S/N)								
Se reconhece, quem é? (escreva o nome da pessoa)								
Grau de certeza no reconhecimento? (Absoluta (A)/ Razoável (R)/ Pequena (P))								

FOTO9	FOTO10	FOTO11	FOTO12

FOTO13	FOTO14	FOTO15	FOTO16

	Foto 9	Foto 10	Foto 11	Foto 12	Foto 13	Foto 14	Foto 15	Foto 16
Reconhece a pessoa? (S/N)								
Se reconhece, quem é? (escreva o nome da pessoa)								
Grau de certeza no reconhecimento? (Absoluta (A)/ Razoável (R)/ Pequena (P))								

FOTO17	FOTO18	FOTO19	FOTO20

FOTO21	FOTO22	FOTO23	FOTO24

	Foto 17	Foto 18	Foto 19	Foto 20	Foto 21	Foto 22	Foto 23	Foto 24
Reconhece a pessoa? (S/N)								
Se reconhece, quem é? (escreva o nome da pessoa)								
Grau de certeza no reconhecimento? (Absoluta (A)/ Razoável (R)/ Pequena (P))								

FOTO25

FOTO26

FOTO27

FOTO28

FOTO29

FOTO30

FOTO31

FOTO32

	Foto 25	Foto 26	Foto 27	Foto 28	Foto 29	Foto 30	Foto 31	Foto 32
Reconhece a pessoa? (S/N)								
Se reconhece, quem é? (escreva o nome da pessoa)								
Grau de certeza no reconhecimento? (Absoluta (A)/ Razoável (R)/ Pequena (P))								

FOTO33	FOTO34	FOTO35	FOTO36

FOTO37	FOTO38	FOTO39	FOTO40

	Foto 33	Foto 34	Foto 35	Foto 36	Foto 37	Foto 38	Foto39	Foto 40
Reconhece a pessoa? (S/N)								
Se reconhece, quem é? (escreva o nome da pessoa)								
Grau de certeza no reconhecimento? (Absoluta (A)/ Razoável (R)/ Pequena (P))								

FOTO41	FOTO42	FOTO43	FOTO44

FOTO45	FOTO46	FOTO47	FOTO48

	Foto 41	Foto 42	Foto 43	Foto 44	Foto 45	Foto 46	Foto 47	Foto 48
Reconhece a pessoa? (S/N)								
Se reconhece, quem é? (escreva o nome da pessoa)								
Grau de certeza no reconhecimento? (Absoluta (A)/ Razoável (R)/ Pequena (P))								

10.5 Texto para projecto fotográfico

Renato Roque: Do retrato enquanto processo de reconhecimento da identidade e dissolução do sujeito

Todos nós já vivemos momentos em que julgamos reconhecer um rosto familiar e nos enganamos. Estamos na rua, num restaurante, num transporte público, e o rosto de alguém associa-se a um rosto que conhecemos, quando, de súbito, concluímos que, afinal, essa não era a pessoa que julgávamos poder identificar. Desfeito o engano, esquecemos esse instante em que um conjunto de complexas operações de percepção e associação cognitiva nos confrontou com as semelhanças e as diferenças detectáveis entre o rosto desconhecido e esse retrato que pensáramos reconhecer dentro da biblioteca mental dos retratos das pessoas nossas conhecidas. Qualquer retrato pressupõe uma identidade ou um processo de identificação. Na história da arte, o retrato foi praticado durante séculos como uma técnica ao serviço do reconhecimento da identidade ou da legitimação de quem tinha o poder para se fazer retratar ou para encomendar o retrato de outra pessoa. De faraós a deuses, dos reis e príncipes a papas e a burgueses abastados, os museus de todo o mundo ilustram uma história de poderes políticos e económicos cuja dimensão simbólica converge nos retratos que conservam. No entanto, os artistas que executavam esses retratos sempre aplicaram também a sua competência técnica de execução a figuras mais ou menos anónimas que não tinham o poder da encomenda, mas que lhes permitiam exercitar a radicalidade de um programa artístico onde a arte sempre se deixou contaminar pela vida, seja a partir da sua representação, seja a partir da sua interpretação e transformação. A par da família real espanhola, Velasquez deixa-nos uma galeria impressionante de retratos dos anões e bufões que povoavam a corte filipina. Flamengos e italianos tinham retratado camponeses nas figuras de santos que acompanhavam frequentes vezes os retratos dos encomendadores de um quadro. O auto-retrato foi igualmente um género praticado como expressão dessa relação com a vida, com as suas mudanças e a sua fugacidade que parecem contradizer a própria natureza de registo que esse auto-retrato poderá documentar. Rembrandt, com os seus auto-retratos realizados ao longo de toda uma vida, será de tal um claro exemplo. Com o aparecimento da fotografia, surge uma real democratização do retrato. Ao mesmo tempo que todos passam a ter a possibilidade de terem o seu retrato, também todos passarão a ser identificáveis por um retrato. Surgem os arquivos de identificação através dos quais todos os países irão cadastrar os seus cidadãos. Estes arquivos são uma evidência do retrato enquanto condição da identidade no seu confronto com a alteridade. Todos são um no conjunto de todos os outros. A similitude do retrato em relação ao retratado é a condição do processo de identificação. No entanto, a arte do século XX, logo a partir dos primeiros Modernismos, nos anos 10 e 20, assumirá como condição da sua revolução estética e conceptual o desmentido da condição aristotélica da representação. Uma grande parte da história da arte do século XX ocupa-se da dissolução do sujeito e da crise da representação que a interrogação da realidade e da condição desse mesmo sujeito poderão pressupor, a partir do momento em que qualquer retrato não será mais do que um espelho superficial, quando conhecida a relevância do inconsciente que a investigação psicanalítica irá revelar. A heteronímia pessoana ou os versos de Mário de Sá-Carneiro "Eu não sou eu nem sou o outro,/Sou qualquer coisa de intermédio" (…) são exemplo dessa dissolução do sujeito que caracterizará a arte moderna. Marcel Duchamp compreenderá que a própria fotografia poderá ser a condição de um labirinto de identidades que confundirá a condição de verdade pressuposta pela sua condição de identificação, quando se faz fotografar por Man Ray enquanto Rrose Scélavy, o seu célebre "alter ego" feminino. Toda uma história prolífica do retrato fotográfico vê-se por sua vez radicalmente revolucionada pelo desenvolvimento da informática e das linguagens computacionais. A fotografia assumirá claramente a possibilidade da representação icónica de uma realidade inexistente, provocadora de uma diferente relação com a percepção e a interpretação da realidade que conhecemos. Pretensas fotografias realistas passam a ser representações de realidades artificialmente encenadas e construídas como num filme, como acontece no caso da obra de Jeff Wall ou retocadas através de programas informáticos que as tornam "mais reais do que o real", como acontece no caso da obra de Thomas Ruff . Em "Espelhos Matriciais", Renato Roque, no contexto do seu projecto de investigação para um programa académico de pós-graduação no domínio da engenharia informática, utiliza novas

possibilidades gráficas de composição e decomposição da imagem que o computador lhe oferece. O fotógrafo parte de retratos por ele obtidos num universo académico de estudantes e professores para os decompor em componentes cuja recombinação lhe permitirá obter como ponto de chegada exactamente o mesmo retrato que é o seu ponto de partida. Um tão quanto complexo quanto laborioso processo de computação gráfica permite-lhe a decomposição e a recomposição da imagem de modo a encontrar as correspondências que possibilitam a percepção da identidade. A reconstrução dessa imagem anteriormente dissolvida em componentes cuja associação será o princípio do processo da sua reconstrução especular constituirá o ponto de chegada de todo este processo. O que parece ser o mesmo não o é, assim como o que parece ser diferente tão pouco o é. Uma curiosa antropomorfização dos processos informáticos materializa o exercício da percepção e do reconhecimento visuais que é condição ontológica desse "deus ex machina" que também é o ser humano. Identidade e alteridade confundem-se e redefinem-se neste processo combinatório. Cada imagem recomposta a partir de uma imagem inicial surge como o resultado de um processo análogo àquele que na literatura replica a história bíblica da criação do homem à imagem de deus em cada narrativa de um autómato construído à imagem do homem, seja o Pinóquio de Carlo Collodi, sejam o Frankenstein de Mary Shelley, a criatura de Hoffman, ou os robôs de Karel Capek , aliás o inventor da palavra "robot", e de Isaac Asimov. Renato Roque apresenta a sua investigação em dois formatos: o formato académico de uma tese de pós-graduação e uma exposição que se apresenta pela primeira vez na Faculdade de Engenharia da Universidade do Porto. Na arte do nosso tempo, a exposição ou a obra de arte é muitas vezes a evidência do seu processo de realização, apresentado através de suportes documentais que fazem parte da sua própria natureza. As fotografias e as imagens que Renato Roque nos apresenta são registos documentais de uma narrativa específica que é a sua própria investigação científica. Importa diferenciar a fotografia da imagem? Serão fotografias os resultados da captura de uma imagem realizada por uma máquina accionada pelo fotógrafo e "imagens" os resultados dos processos gráficos realizados pelo computador? Serão ambas imagens de diferente natureza? Como articular a diferença dos seus processos de construção com a semelhança dos seus resultados? Como confrontar as possibilidades de falsificação do real que estes mesmos processos informáticos possibilitam, hoje facilmente praticáveis através dos vários programas de "morphing" existentes? Ciência e arte cruzam-se nos processos técnicos que materializam estas imagens e que nos propõem um confronto com a percepção dos retratos que reproduzem, descaracterizando-os e recaracterizando-os em operações que nos confrontam com a própria natureza da percepção inerente a qualquer operação de identificação visual. A redundância é extrema, a ponto de nos deixar a interrogação do porquê de uma investigação que replica um retrato num outro retrato exactamente igual. No entanto, essa redundância torna-se paradoxal quando confrontada com a diversidade das imagens que encontramos nas componentes obtidas de cada retrato. Extrair o universal do particular sempre foi característico dos processos científicos. Indistinguir o particular do universal sempre foi característico dos processos de criação artística. Em ambos os processos, a finalidade ou o objectivo prático da construção da percepção ou do conhecimento são matérias irrelevantes. Esta é a ambiguidade que Renato Roque nos oferece, a mesma ambiguidade geradora da obra de arte que fará parte da nossa natureza humana (Joseph Beuys assumia que cada homem era um artista…) e que assumimos no dia a dia, entre a percepção e o conhecimento, como em cada momento desse conflito entre o conhecido e o desconhecido que sucede quando pensamos reconhecer um rosto na multidão que se vem a revelar muitas vezes a expressão de uma decepção das nossas expectativas prévias ou a reconfiguração do desejo de outros rostos que ambicionaríamos reconhecer.

João Fernandes, Julho 2009

10.6 Publicações realizadas

Até à realização das provas de mestrado referentes a esta tese de dissertação foram aceites e publicados os seguintes artigos:

- Artigo "Podemos construir retratos usando um retratário?", publicado no **CIANTEC** 2009 na UA

- Artigo "On building photographic portraits using statistical techniques ", publicado no **VIP IMAGE 2009** na FEUP