

**FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO**

# **Impacto dos Filtros de Abstracção no Reconhecimento Facial em Imagens**

**Pedro Tiago Pontes**

**PREPARAÇÃO DA DISSERTAÇÃO**



Mestrado Integrado em Engenharia Informática e Computação

Orientador: Cristina Ribeiro

Co-orientador: Filipe Coelho

06 de Fevereiro de 2013

© Pedro Tiago Pontes, 2013

# **Impacto dos Filtros de Abstração no Reconhecimento Facial em Imagens**

**Pedro Tiago Pontes**

Mestrado Integrado em Engenharia Informática e Computação

# Resumo

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam vitae quam sed mauris auctor porttitor. Mauris porta sem vitae arcu sagittis facilisis. Proin sodales risus sit amet arcu. Quisque eu pede eu elit pulvinar porttitor. Maecenas dignissim tincidunt dui. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec non augue sit amet nulla gravida rutrum. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Nunc at nunc. Etiam egestas.

Donec malesuada pede eget nunc. Fusce porttitor felis eget mi mattis vestibulum. Pellentesque faucibus. Cras adipiscing dolor quis mi. Quisque sagittis, justo sed dapibus pharetra, lectus velit tincidunt eros, ac fermentum nulla velit vel sapien. Vestibulum sem mauris, hendrerit non, feugiat ac, varius ornare, lectus. Praesent urna tellus, euismod in, hendrerit sit amet, pretium vitae, nisi. Proin nisl sem, ultrices eget, faucibus a, feugiat non, purus. Etiam mi tortor, convallis quis, pharetra ut, consectetuer eu, orci. Vivamus aliquet. Aenean mollis fringilla erat. Vivamus mollis, purus at pellentesque faucibus, sapien lorem eleifend quam, mollis luctus mi purus in dui. Maecenas volutpat mauris eu lectus. Morbi vel risus et dolor bibendum malesuada. Donec feugiat tristique erat. Nam porta auctor mi. Nulla purus. Nam aliquam.

# Abstract

Here goes the abstract written in English.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed vehicula lorem commodo dui. Fusce mollis feugiat elit. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Donec eu quam. Aenean consectetur odio quis nisi. Fusce molestie metus sed neque. Praesent nulla. Donec quis urna. Pellentesque hendrerit vulputate nunc. Donec id eros et leo ullamcorper placerat. Curabitur aliquam tellus et diam.

Ut tortor. Morbi eget elit. Maecenas nec risus. Sed ultricies. Sed scelerisque libero faucibus sem. Nullam molestie leo quis tellus. Donec ipsum. Nulla lobortis purus pharetra turpis. Nulla laoreet, arcu nec hendrerit vulputate, tortor elit eleifend turpis, et aliquam leo metus in dolor. Praesent sed nulla. Mauris ac augue. Cras ac orci. Etiam sed urna eget nulla sodales venenatis. Donec faucibus ante eget dui. Nam magna. Suspendisse sollicitudin est et mi.

Fusce sed ipsum vel velit imperdiet dictum. Sed nisi purus, dapibus ut, iaculis ac, placerat id, purus. Integer aliquet elementum libero. Phasellus facilisis leo eget elit. Nullam nisi magna, ornare at, aliquet et, porta id, odio. Sed volutpat tellus consectetur ligula. Phasellus turpis augue, malesuada et, placerat fringilla, ornare nec, eros. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Vivamus ornare quam nec sem mattis vulputate. Nullam porta, diam nec porta mollis, orci leo condimentum sapien, quis venenatis mi dolor a metus. Nullam mollis. Aenean metus massa, pellentesque sit amet, sagittis eget, tincidunt in, arcu. Vestibulum porta laoreet tortor. Nullam mollis elit nec justo. In nulla ligula, pellentesque sit amet, consequat sed, faucibus id, velit. Fusce purus. Quisque sagittis urna at quam. Ut eu lacus. Maecenas tortor nibh, ultricies nec, vestibulum varius, egestas id, sapien.

Donec hendrerit. Vivamus suscipit egestas nibh. In ornare leo ut massa. Donec nisi nisl, dignissim quis, faucibus a, bibendum ac, diam. Nam adipiscing hendrerit mi. Morbi ac nulla. Nullam id est ac nisi consectetur commodo. Pellentesque aliquam massa sit amet tellus. Vivamus sodales aliquam leo.

# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Contexto . . . . .	1
1.2	Motivação . . . . .	2
1.3	Objetivos . . . . .	2
1.4	Estrutura do Relatório . . . . .	2
<b>2</b>	<b>Reconhecimento Facial em Imagens</b>	<b>3</b>
2.1	Definição do Problema . . . . .	4
2.1.1	Sub-Problemas . . . . .	5
2.1.2	Desafios . . . . .	6
2.2	Áreas de Aplicação . . . . .	8
2.2.1	Controlo de Acesso e Segurança de Informação . . . . .	8
2.2.2	Cartões inteligentes e Identificação de Faces . . . . .	9
2.2.3	Entretenimento . . . . .	9
2.2.4	Gestão de Conteúdos Multimédia e Bases de Dados . . . . .	10
2.2.5	Segurança e Reforço da Lei . . . . .	11
2.3	Estratégias de Reconhecimento Facial . . . . .	11
2.3.1	Soluções Existentes . . . . .	13
2.4	Evolução Desempenho . . . . .	14
<b>3</b>	<b>Abstração de Imagens</b>	<b>20</b>
3.1	Filtros de Abstração de Imagens . . . . .	21
3.1.1	Filtro Bilateral . . . . .	21
3.1.2	Kuwahara . . . . .	22
3.1.3	<i>Diffusion Shock Filter</i> . . . . .	23
3.1.4	Filtragem Morfológica . . . . .	24
3.1.5	Técnicas <i>Gradient Domain</i> . . . . .	24
3.2	Trabalho Preliminar . . . . .	24
<b>4</b>	<b>Perspectiva de Solução</b>	<b>25</b>
4.1	Objetivos . . . . .	26
4.2	Implementação . . . . .	26
4.2.1	Sistema de Reconhecimento Facial Base - OpenCV Face Recognizer . . . . .	27
4.2.2	Filtros de Abstração - Filtro <i>Anisotropic Kuwahara</i> . . . . .	28
4.2.3	Coleções de dados . . . . .	29
4.3	Avaliação Desempenho . . . . .	30
4.3.1	Identificação . . . . .	32
4.3.2	Verificação . . . . .	33

## CONTEÚDO

4.4 Resultados Esperados . . . . .	34
4.5 Plano de Trabalho . . . . .	35
<b>5 Conclusão</b>	<b>37</b>
<b>Referências</b>	<b>38</b>

# **Lista de Figuras**

2.1	Sistema Genérico de Reconhecimento Facial . . . . .	6
2.2	A redução da taxa de erro para os algoritmos estados da arte tal como documentado nas avaliações FERET, FRVT 2002, FRVT 2006 e MBE 2010 [PSO07, GQP10] .	18
3.1	Cronologia do desenvolvimento da área da IB-AR [KCWI12]. . . . .	20
3.2	Diferentes filtros de abstração aplicados sobre a mesma imagem [KKD09]. . . . .	22
3.3	<i>Diffusion Shock Filter</i> [KL08]. . . . .	23
4.1	Comparação entre imagem não abstraída (esquerda) e imagem abstraída (direita).	29
4.2	Exemplos de imagens da coleção Sapo Fama. . . . .	30
4.3	Galeria com faces conhecidas pelo sistema e respectivas anotações. . . . .	31
4.4	Imagens de pessoas conhecidas pelo sistema, mas não presentes na galeria. . . . .	32
4.5	Imagens de pessoas não conhecidas pelo sistema. . . . .	33
4.6	ROC. . . . .	34
4.7	ROC2. . . . .	35
4.8	Plano de Trabalho 1º e 2º Semestres . . . . .	36

# **Lista de Tabelas**

2.1 Categorização dos Métodos de Reconhecimento Facial em Imagens por Zhao <i>et al.</i> [ZCPR03] . . . . .	12
-------------------------------------------------------------------------------------------------------------	----

# Abreviaturas e Símbolos

UNCOL	UNiversal COmpiler-oriented Language
Loren	Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Sed vehicula lorem commodo dui
WWW	<i>World Wide Web</i>

# Capítulo 1

## Introdução

TO-DO NOTA: Versão antiga feita perto da antiga apresentação

### 1.1 Contexto

A sociedade atual caracteriza-se pelo constante fluxo de informação a que os seus indivíduos estão expostos. A título de exemplo e de acordo com dados divulgados pela Intel em [Int], a cada minuto, na Internet são visualizados no Youtube 1,3 milhões de vídeos e carregadas para os seus servidores 30 novas horas de vídeo. Já o Flickr, no mesmo período de tempo, regista a visualização de 20 milhões de fotos e o carregamento de 3000 novas imagens para a sua plataforma. Segundo os mesmos dados, atualmente o número de dispositivos com ligação à Internet é equivalente ao da população mundial e é esperado que em 2015 seja o dobro dessa população. Este aumento e diversificação de conteúdos digitais e respetivos dispositivos que a eles acedem, torna particularmente difícil a sua categorização, organização e posterior pesquisa por parte dos utilizadores. A área da Recuperação de Informação Multimédia (RIM), que trata a pesquisa de informação em todas as suas formas, ganha assim especial importância na atualidade, nomeadamente através da utilização de métodos baseados em conteúdo na pesquisa de informação, os quais tem a potencialidade de melhorar a exatidão da informação extraída, mesmo quando existem anotações textuais acerca dos conteúdos, ao possibilitar uma visão adicional sobre as coleções multimédia.[LSDJ06].

O reconhecimento facial é um componente importante da capacidade de percepção humana, sendo efetuado de uma forma rotineira e quase sem esforço por parte dos seres humanos. Contudo, a construção de sistemas computacionais capazes de efetuar este tipo de reconhecimento de uma forma semelhante aos humanos em termos de eficácia é ainda uma área de investigação em aberto. Este é um problema de tal forma desafiante e interessante que tem despertado a atenção de investigadores das mais diversas áreas, ao longo dos últimos 40 anos, tais como: psicologia, reconhecimento de padrões, redes neurais, visão por computador e computação gráfica. [ZCPR03].

## Introdução

De uma forma geral o problema de reconhecimento facial pode ser formulado da seguinte forma: dada uma imagem pretende-se identificar ou verificar a presença de uma ou mais pessoas na cena, comparando-a com uma base de dados de faces previamente guardada. A solução deste problema envolve três passos fundamentais: Deteção e segmentação das faces presentes na imagem, extração dos atributos faciais e por fim identificação ou verificação da identidade da pessoa.[[ZCPR03](#)]

### 1.2 Motivação

A pesquisa ao nível do reconhecimento facial é motivada não só pelos desafios inerentes ao processo de reconhecimento facial, mas também pelos inúmeras aplicações onde a identificação de indivíduos é necessária [[LJ11](#)]. Esta é uma área de investigação em aberto e para a qual ainda existe uma grande margem de progressão, no entanto, tal como mencionado anteriormente, esta é uma tarefa que envolve um conjunto grande de subproblemas que devem ser abordados separadamente.

No âmbito desta dissertação o impacto de filtros visuais de abstração de informação no processo de reconhecimento facial será o principal foque de investigação. Este filtros permitem uma simplificação do conteúdo visual ao retirar informação redundante e destacando a apenas a informação essencial. Estes filtros são tradicionalmente utilizados por artistas para comunicar de uma forma mais eficaz a informação pretendida ao retirar a informação desnecessária de imagens foto-realistas [[KKD09](#)]. Estudos efetuados demonstraram que aplicação destes filtros no âmbito da abstração de imagens para a ilustração automática de texto têm a potencialidade de melhorar a informação retornada, assim como reduzir significativamente as necessidades de processamento e armazenamento [?]. Tendo em conta os resultados previamente obtidos, e uma vez que não existem estudos relativos à utilização destes filtros no processo de reconhecimento facial, torna-se pertinente o estudo do seu impacto no âmbito desta dissertação.

Falar da crescente importância na atualidade e porquê - aplicações comerciais

### 1.3 Objetivos

O principal objetivos deste projeto visa o estudo do impacto de filtros visuais de informação de contexto no processo de reconhecimento facial. Adicionalmente propõe-se também o desenvolvimento de um sistema de reconhecimento facial de personalidades que integre a abstração de imagens no processo de reconhecimento facial. Este sistema deverá ser capaz de detetar as faces presentes numa imagem e apresentar uma lista de possíveis entidades nela contidas.

### 1.4 Estrutura do Relatório

## Capítulo 2

# Reconhecimento Facial em Imagens

A capacidade de distinguir indivíduos através das suas características faciais é inata aos seres humanos, efetuada uma forma natural e quase sem esforço. A forma como efetuamos esse reconhecimento, é no entanto um problema complexo, que atrai a atenção de investigadores das mais diversas áreas, sendo da psicologia e na década de 50 [BT54], o mais antigo estudo efetuado sobre o tema. Associado a este problema, o reconhecimento facial automático, ou seja, o reconhecimento facial efetuado de uma forma automática com o auxílio de máquinas tais como computadores, tem de igual forma recebido muita atenção por parte da comunidade científica, sendo que o primeiro sistema de reconhecimento facial automático de que há registo, foi desenvolvido da década de 70 por Takeo Kanade [Kan73]. Ao longo dos anos, investigadores das mais diversas áreas, tais como psicologia, reconhecimento de padrões, neurologia, visão por computador ou computação gráfica, efetuaram estudos em vários aspectos do reconhecimento facial efetuado por computadores e máquinas.

Ao nível do reconhecimento facial automático, após o trabalho de Kanade, e outros realizados por alguns dos seus pares na década de 70, verificou-se um abrandamento na investigação científica efetuada na área até meados dos anos 90, onde o trabalho realizado por Turk e Pentland [TP91] trouxe um novo impulso ao desenvolvimento de sistemas e metodologias de reconhecimento facial automático. Neste estudo, Turk e Pentland apresentaram uma nova forma abordagem ao reconhecimento facial designada de *Eingenface*. Um outro marco no desenvolvimento nas metodologias de reconhecimento facial, foi o surgimento do método *Fisherface* [BHK97, ZCK98], o qual apresenta uma maior insensibilidade em relação à variação da iluminação e expressões faciais presentes nas imagens.

Ao longo da década de 90 a investigação científica na área do reconhecimento facial automático intensificou-se, tendência que ainda se verifica atualmente, tal como pode ser verificado pelo surgimento de várias conferências, das quais são exemplo, desde as mais antigas *International Conference on Automatic Face and Gesture Recognition (AFGR)* (1995) e *International Conference on Audio- and Video-Based Authentication (AVBPA)* (1997), até à recentemente realizada *5th International Conference on Biometrics (ICBS)* (2012).

Paralelamente, o surgimento de programas como o Face Recognition Technology (FERET)

(1993-1998), levado a cabo pelo *National Institute of Standards and Technology (NIST)* dos E.U.A, com o objetivo de financiar a investigação na área e estabelecer padrões para a avaliação das diversas metodologias de reconhecimento facial criadas, assim como impulsionar a criação de uma base de dados de faces de grande dimensões para o teste dos sistemas criados, permitiu uma evolução significativa na área [[PWHR98](#), [PRR00](#)].

No final da década de 90, surgiram as primeiras aplicações comerciais de reconhecimento facial automático, tal como pode ser comprovado pela participação de 5 soluções comerciais na primeira edição dos *Facial Recognition Vendor Test (FRVT)*, no ano 2000 [[BBP01](#)]. Nesta avaliação, as soluções eram testadas nas componentes de usabilidade e performance do reconhecimento efetuado, de modo fornecer uma avaliação objetiva e independente dos sistemas de reconhecimento facial presentes no mercado. Estas avaliações foram repetidas nos anos de 2002, 2006, 2010 e 2012.

Neste capítulo após uma breve introdução ao problema de reconhecimento facial automático, é apresentada uma abordagem genérica para a solução do problema. Em seguida, são apresentadas as diversas áreas de aplicação do reconhecimento facial automático. Por último, é apresentada uma revisão do estado da arte dos sistemas atuais e uma análise da evolução da sua performance com base nas avaliações efetuadas no âmbito dos programas FERET e FRVT.

## 2.1 Definição do Problema

De uma forma geral, o problema de reconhecimento facial pode ser formulado da seguinte forma: dada uma imagem (prova), pretende-se identificar ou verificar a presença de uma ou mais pessoas na prova, comparando-a com uma base de dados de faces previamente guardada (galeria) [[YKMA02](#), [ZCP03](#)]. Este problema, pode ainda ser dividido em dois sub-problemas específicos: verificação de faces (ou autenticação) e identificação de faces [[LJ11](#)]:

**Verificação** Uma prova é apresentada ao sistema, assim como a respetiva identidade da pessoa presente na imagem. Cabe ao sistema verificar se existe uma correspondência entre a pessoa representada na imagem e a identidade fornecida. Assim, é efetuada uma comparação 1:1, que responde à pergunta: "Esta pessoa é quem realmente afirma ser?".

**Identificação** Uma prova é apresentada ao sistema, e pretende-se que este devolva a identidade da pessoa presente na imagem, assim como o nível de confiança associado à correspondência efetuada entre a imagem e a identidade devolvida. A identificação de faces pode ainda ser dividida em dois sub-problemas:

No primeiro problema designado *watch list* ou *open-set identification*[[CSP10](#)], não é garantido que o sistema conheça a pessoa a ser identificada, pelo que em primeiro lugar é necessário identificar se a pessoa existe na galeria e, caso exista, determinar a identidade da pessoa. Este sub-problema responde às perguntas: "Esta pessoa está presente na galeria? Se sim, de quem é esta face?".

O segundo caso, é aquele que é geralmente avaliado em sistemas de reconhecimento facial, onde é garantido que a pessoa é conhecida pelo sistema, respondendo-se apenas à pergunta: "De quem é esta face?". Nesta situação, designada de *closed-set identification*, podem ainda ser pedidas todas as correspondências efetuadas com um nível de confiança superior a um valor pré-definido. Em qualquer um dos subproblemas de identificação, é efetuada uma comparação 1:N.

Os sistemas de reconhecimento facial podem ainda ser divididos genericamente em dois grupos, tendo em conta a fonte de onde são extraídas as faces a analisar. O primeiro grupo baseia-se no uso de imagens estáticas, como por exemplo fotografias de personalidades ou mesmo em retratos de faces humanas criadas por ilustradores. O segundo, tem por base o uso de vídeos, dos quais são extraídas as faces a identificar ou verificar, sendo que neste caso essa identificação pode ainda ser efetuada em tempo real ou *à posteriori*. No entanto, cada um destes meios representa problemas e desafios diferentes ao reconhecimento facial automático, pelo que devem ser estudados de forma individualizada, sendo que, no âmbito desta dissertação irá apenas ser abordado o sub-problema de reconhecimento facial efetuado em imagens estáticas.

### 2.1.1 Sub-Problemas

O problema de reconhecimento facial em imagens é um problema complexo que obriga a que um conjunto de sub-tarefas sejam efetuadas com sucesso para uma identificação ou verificação eficaz da pessoa presente na imagem. Em primeiro lugar, é necessário determinar onde se encontra a face da pessoa a ser identificada, separando-a do resto da imagem. Posteriormente é necessário encontrar características que diferenciem essa face de outras. Finalmente, é ainda necessário efetuar a comparação da face, e respetivas características, com a galeria de imagens existente, fazendo assim o reconhecimento pessoa representada. Este conjunto de sub-tarefas, são geralmente designados pelos seguintes três passos fundamentais [ZCPR03]:

- 1. Detecção** Determina a presença de uma face na imagem, e em caso positivo, efetua a segmentação do resto da imagem.
- 2. Extração características faciais** Consiste na extração de características úteis para a diferenciação entre as diversas faces presentes no sistema. As características extraídas variam dependendo do método utilizado, podendo ir desde a distância entre olhos, à textura e cor da pele da pessoa a ser reconhecida, por exemplo.
- 3. Identificação ou Verificação** Consiste na identificação ou verificação da identidade do sujeito presente na imagem, tendo em conta as características extraídas e a informação presente no sistema.

Sendo que, um quarto passo, efetuado após a deteção e antes da extração das características faciais, pode ainda ser destacado [LJ11]:

**Normalização** Introduzida em alguns métodos de reconhecimento facial com o objetivo de melhorar os resultados do reconhecimento ao nível pose e iluminação das faces nas imagens. O seu objetivo consiste em normalizar a face geometricamente, ou seja na sua forma e enquadramento, e fotometricamente, ou seja ao nível da iluminação.

Na figura 2.1 é possível ver uma representação dos quatro passos fundamentais enunciados anteriormente que constituem um sistema de reconhecimento facial genérico. Em algumas soluções os passos de deteção e extração de características possuem tarefas em comum, sendo realizados simultaneamente.

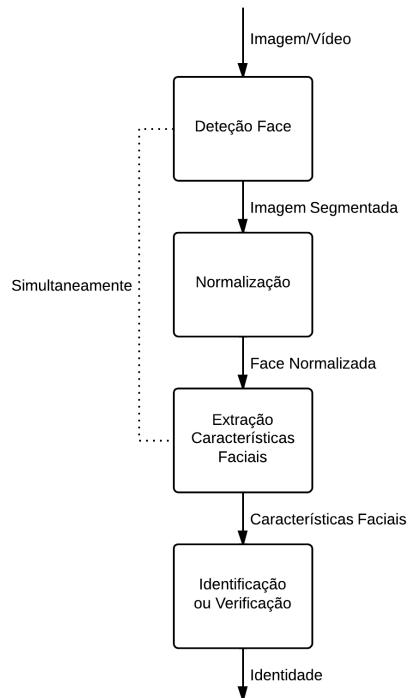


Figura 2.1: Sistema Genérico de Reconhecimento Facial

### 2.1.2 Desafios

Sendo um problema complexo, existem um conjunto de desafios que se colocam na construção de um sistema de reconhecimento facial automático.

O primeiro passo em qualquer um desses sistemas, é a deteção de faces na imagem, a esse nível podem destacar-se os seguintes desafios [YKMA02]:

1. Pose e Orientação: A posição relativa da face à câmara pode sofrer uma grande variação dependendo do ambiente onde é capturada a imagem (variação no ângulo da face em relação à câmara vertical e horizontalmente, posição frontal, perfil). Ao variar a posição da

## Reconhecimento Facial em Imagens

face, alguns dos elementos que a permitem localizar e reconhecer podem ficar obstruídos. Avaliações efetuadas [BBP01], demonstram que este impacto é maior, quanto maior for o ângulo da pose em relação à posição frontal, sendo que uma variação até cerca de 25º não afeta significativamente os resultados do reconhecimento.

2. Elementos Estruturais: a presença ou ausência de elementos estruturais, tais como barba ou óculos, assim como, a grande variação de forma e cor a que estes elementos podem apresentar afetam significativamente a forma como a face é percecionada pelo sistema.
3. Obstrução da face: as imagens captadas podem possuir objetos em frente à cara da pessoa a ser identificada, levando apenas a uma representação parcial da face na imagem.
4. Condições da imagem: aspectos como a resolução da imagem, o sensor utilizado para a captura e até mesmo eventuais defeitos associados à câmara utilizada podem potencialmente afetar o aspeto final da face na imagem.

Outro fator que afeta significativamente a percepção de uma face por um sistema de reconhecimento facial automático é a iluminação. Uma imagem pode ser capturada, com iluminação natural ou artificial. Para além disso, a iluminação pode estar sujeita a uma grande variação de tonalidades, intensidade e ângulo de incidência da luz.

Mesmo quando captadas pela mesma câmara, com as mesmas condições de iluminação e no mesmo local, podem existir variações significativas nas imagens captadas da face de uma pessoa devido às diferentes expressões faciais contidas nas imagens, fazendo com que este seja outro desafio a ultrapassar ao efetuar o reconhecimento facial automático.

O envelhecimento, ou eventuais alterações provocadas na face devido a fatores externos como por exemplo acidentes, exposição a condições extremas ou intervenção cirúrgica é outro dos fatores a ter em conta na identificação de uma pessoa.

Finalmente, existem ainda fatores como o elevado grau de semelhança entre pessoas com graus de parentesco próximos. Imagens capturadas de gémeos, ou até mesmo de pais e filhos, podem à primeira vista parecer imagens da mesma pessoa, tornando ainda mais difícil para um sistema automático efetuar a distinção dos indivíduos.

Assim sendo, para um reconhecimento fiável e eficaz é essencial que todas as etapas identificadas em ?? sejam concluídas com sucesso. Sem uma correta deteção da presença de uma face na imagem, todos os passos subjacentes seriam impossíveis de ser realizados, por outro lado, o sucesso da extração das características faciais, encontram-se também dependente de uma normalização eficaz das imagens. No entanto, apesar desta dependência verificada entre as diversas etapas, cada uma delas representa problemas complexos, que devem merecer a atenção e estudo de forma individual para que uma maior evolução das diversas técnicas envolvidas seja atingida [ZCPR03].

## 2.2 Áreas de Aplicação

O problema de reconhecimento facial automático tem merecido a atenção e estudo de inúmeros investigadores ao longo dos últimos 40 anos, motivados não só pelos desafios inerentes ao processo de reconhecimento facial, mas também pelas vasta gama aplicações onde a identificação de indivíduos é necessária [LJ11].

Adicionalmente, a evolução tecnológica da última década, permitiu a criação de sistemas computacionais mais poderosos, os quais tornaram possíveis desenvolvimentos na área que anteriormente seriam impensáveis gerando um ainda maior interesse, tal como pode ser verificado pelo surgimento de uma vasta gama de aplicações comerciais e governamentais de reconhecimento facial, como é exemplo o sistema de fronteira automática dos aeroportos portugueses [Minb], assim como, pela recente aquisição de soluções de reconhecimento facial [?, ?] por parte de duas grandes empresas tecnológicas nomeadamente Google e Facebook.

De seguida encontram-se descritas, em 6 categorias, algumas das áreas de aplicação do reconhecimento facial ilustradas com exemplos de aplicação real. Esta categorização foi baseada no estudo efetuado em [LJ11], o qual é para nosso conhecimento o mais recente estudo efetuado às diversas áreas de aplicação do reconhecimento facial, assim como em estudos anteriores efetuados por [ZCPR03]. Esta secção pretende apenas apresentar de uma forma sumária algumas das possíveis áreas de aplicação e respetivas soluções às quais o reconhecimento facial poderá ser aplicado, não pretendendo ser um estudo exaustivo de todas as soluções existentes, uma vez que isso sairia do âmbito desta dissertação.

### 2.2.1 Controlo de Acesso e Segurança de Informação

A necessidade de sistemas de fácil utilização que assegurem a segurança da nossa informação digital, assim como da privacidade dos utilizadores, sem que para isso seja necessário a memorização e utilização de um sem número de palavras-chave e códigos de segurança é premente. Neste campo, o reconhecimento facial destaca-se ao permitir a criação de sistemas de autenticação não invasivos que requerem um nível de cooperação baixo por parte dos participantes, em contraste com os sistemas tradicionais como o uso de palavras-chave, ou mesmo de sistemas mais complexos como a análise de impressões digitais, retina ou íris. Por outro lado, este tipo de sistemas, permite ainda determinar a presença factual do utilizador que se pretende identificar, e não apenas determinar o conhecimento de um código de acesso à informação.

Atualmente, o uso de reconhecimento facial para o controlo de acesso a informação encontra-se já a ser utilizado em diversas aplicações comerciais como são exemplos os sistemas de autenticação automática disponíveis numa grande variedade de computadores portáteis ???. De uma forma geral, nestes sistemas, o utilizador apenas necessita de se aproximar do computador para que o reconhecimento facial e consequente autenticação sejam efetuados. Desta forma o utilizador não necessita de efetuar qualquer intervenção para aceder aos seus dados ou de memorizar qualquer palavra-chave.

O uso de reconhecimento facial para garantir uma maior segurança de informação, ou efectuar um mais eficaz controlo do seu acesso, é geralmente aplicado em soluções com um número de utilizadores limitado e nas quais as imagens capturadas apresentam baixa variabilidade nas condições de iluminação e pose dos utilizadores. Desta forma, o nível de precisão do reconhecimento atingido nesta área é normalmente elevado, fazendo com que o nível de satisfação dos seus utilizadores seja também elevado [LJ11].

### 2.2.2 Cartões inteligentes e Identificação de Faces

Os documentos pessoais inteligentes registam uma presença cada vez mais ativa na sociedade atual, como é caso do cartão do cidadão português (CC), que à data de 5 de Novembro de 2012 já tinha sido adotado por mais de 7 milhões de portugueses [Age], ou mesmo do passaporte eletrónico português (PEP), o qual recebeceu igualmente uma ampla aceitação desde a sua criação em 2006 [Mina].

Tradicionalmente este tipo de documentos encontram-se dotados de capacidade de processamento e armazenamento próprios, podendo interagir com aplicações de terceiros como, por exemplo, sistemas de autenticação com recurso ao reconhecimento facial. O sistema designado de Reconhecimento Automático de Passageiros Identificados Documentalmente (RAPID)[Minb], desenvolvido pela empresa portuguesa Vision-Box [VB] encontra-se desde 2007 em funcionamento nos vários aeroportos internacionais portugueses e tira partido das vantagens da associação da tecnologia de reconhecimento facial com documentos inteligentes. Em primeiro lugar, o sistema verifica se os dados armazenados no chip do passaporte são válidos, posteriormente, é efetuada uma comparação entre a fotografia armazenada no cartão e a imagem capturada do passageiro pelo dispositivo instalado na fronteira automática. Caso a identificação seja verificada, a porta que permite a passagem entre fronteiras é então automaticamente aberta. Todo o processo, encontra-se ainda a ser supervisionado por um operador numa cabine, o qual se encontra a supervisionar múltiplos dispositivos de identificação, sendo apenas necessária a sua intervenção quando não é efetuada uma identificação válida da pessoa.

O nível de precisão dos sistemas que tiram partido das sinergias entre cartões inteligentes e o reconhecimento facial apresenta uma correlação com o nível de atualização dos dados presentes nos documentos, sendo que com um nível de atualização frequente é possível ser atingido um nível de satisfação elevado [LJ11], como se comprova pela bom funcionamento do sistema RAPID. Contudo, esta é ainda uma área de investigação em aberto para que uma utilização destes sistemas de forma totalmente autónoma (sem que seja necessária a supervisão de um operador), em larga escala e, por exemplo, em ambientes não controlados seja possível.

### 2.2.3 Entretenimento

As tecnologias de reconhecimento facial apresentam uma forte potencialidade de aplicação nas áreas associadas ao entretenimento e à produção de conteúdos digitais, nomeadamente no que diz respeito à produção de jogo de computador e interação humano computador.

As consolas modernas estão dotadas de capacidade de processamento notáveis, para além disso, juntamente com estas consolas é cada vez mais comum o surgimento de dispositivos como o *Kinect* da Microsoft que permitem novas formas de interação e a criação de novos tipos de jogos. Uma das novidades introduzidas pelo Kinect é a inclusão de diversos sensores e câmaras que permitem, entre outras coisas, efetuar um reconhecimento facial dos utilizadores. Desta forma, as equipas de desenvolvimento de jogos de computador, podem agora de uma forma mais fácil utilizar mecanismos de reconhecimento facial automático nos seus jogos de forma a criar jogos mais imersivos e personalizados. Este reconhecimento pode ainda ser aplicado ao nível da interação pessoa-computador onde permite, a criação de sistemas com uma maior facilidade de interação, uma vez que o próprio rosto poderá servir como um dispositivo de *input* natural.

Á semelhança dos sistemas de controlo de acesso o reconhecimento facial automático na área do entretenimento é geralmente aplicado em soluções com um número de utilizadores limitado e onde o nível de precisão necessário não é muito alto fazendo assim com que a satisfação dos seus utilizadores seja, em geral, elevada.

### 2.2.4 Gestão de Conteúdos Multimédia e Bases de Dados

A sociedade atual caracteriza-se pelo constante fluxo de informação a que os seus indivíduos se encontram expostos. A título de exemplo e de acordo com dados divulgados pela Intel[[Int](#)], a cada minuto, são visualizados no Youtube 1,3 milhões de vídeos e carregadas para os seus servidores 30 novas horas de vídeo. Já o Flickr, no mesmo período de tempo, regista a visualização de 20 milhões de fotos e o carregamento de 3000 novas imagens para a sua plataforma. Segundo os mesmos dados, atualmente o número de dispositivos com ligação à Internet é equivalente ao da população mundial e é esperado que em 2015 seja o dobro dessa população.

Este crescimento verificado quer na quantidade de conteúdos produzidos, quer na diversidade de dispositivos que acedem a esses conteúdos, torna assim extremamente importante uma gestão e organização eficazes dos mesmos. Os métodos tradicionais de pesquisa e organização de conteúdos multimédia baseiam-se em anotações textuais associadas a estes conteúdos designadas de descritores. Com o uso de sistemas de reconhecimento facial é possível efetuar o reconhecimento das entidades presentes numa biblioteca multimédia mesmo que as suas fotografias não possuam qualquer anotação textual, uma vez que é feita uma análise da própria imagem e não apenas dos seus descritores, tornando-se assim um importante auxílio dos métodos tradicionais.

Atualmente existem no mercado diversas aplicações que permitem efetuar a organização de álbuns fotográficos pessoais com recurso ao reconhecimento facial como por exemplo o Picasa da Google ou o iPhoto da Apple. Por outro lado, redes sociais como Google Plus e o Facebook, lançaram também recentemente serviços que permitem aos seus utilizadores a deteção e identificação automática de outros utilizadores presentes nas imagens carregadas para os seus servidores.

Apesar de a performance destes sistemas ter registado uma evolução notável nos últimos anos, a grande variabilidade associada às condições de captura de imagens e a grande quantidade de dados a ser processada faz com que exista ainda uma grande margem de evolução possível a este

nível, sendo que a performance verificada ao nível do reconhecimento efetuado varia muito de acordo com a aplicação a analisar, assim como as características do conjunto de dados a organizar.

### 2.2.5 Segurança e Reforço da Lei

A preocupação com a segurança de instalações públicas e privadas, torna-se ainda mais premente em situações de grave crise económica como aquela que vivemos atualmente. Por outro lado, fatores como o controlo de emigração ilegal, preocupações relacionadas com a segurança de eventos com um grande impacto mediático ou mesmo o auxílio à investigação policial tornaram esta uma das áreas de aplicação que mais interesse despoletou por parte de sectores privados e governamentais desde o surgimento das primeiras tecnologias de reconhecimento facial.

Uma das aplicações mais comuns do reconhecimento facial em segurança é o reforço da vigilância da via pública. Desde 1998, diversas cidades implementaram sistemas de vigilância com recurso a reconhecimento facial, como são exemplos os sistemas instalados em *Newham Borough of London*, *Tampa (Florida)* e *Virginia Beach (Virginia)* [LJ11]. Um outro exemplo, foi o sistema desenvolvido para os jogos olímpicos de 2008 em Pequim pela *Chinese Academy Of Sciences* e que foi utilizado para validar as entradas nas cerimónias de abertura e encerramento dos jogos olímpicos [Chi].

A aplicação de sistemas de reconhecimento facial para a segurança e reforço da lei tem atraído muita atenção por parte de sectores privados e governamentais, havendo o registo de alguns casos onde a sua aplicação se revelou um sucesso. Contudo, de uma forma geral estes sistemas debatem-se com um baixo nível de satisfação por parte dos seus utilizadores, isto porque, existe uma grande variabilidade nas condições de iluminação das imagens captadas, pose dos utilizadores, assim como, um número muito elevado de faces a analisar, resultando num número elevado de falsos positivos e consequentemente numa má performance dos sistemas [LJ11].

## 2.3 Estratégias de Reconhecimento Facial

Após o trabalho inicial de Takeo Kanade em 1973 [Kan73], a área do reconhecimento facial automático passou um período de dormência até meados da década de 90, onde os trabalhos realizados recorrendo a métodos de *still-face-recognition subspace* tais como *Principal Component Analysis (PCA)*, *Linear Discriminant analysis (LDA)* e *Elastic Graph Matching (EGM)*, revitalizaram a área [CSP10]. Ao nível da PCA, o método *Eingenface* [TP91] desenvolvido por Turk e Pentland em 1991, baseado na representação facial de baixo nível com base na análise dos componentes principais das representações faciais desenvolvida anteriormente por Kirby e Sirovich [KS90], foi um marco fundamental no rejuvenescimento do reconhecimento facial automático. Posteriormente, o método *Fisherfaces* [BHK97, EC97, ZCK98], que tira partido da aplicação de LDA após uma análise inicial dos componentes principais da imagem, apresentou também resultados positivos em experiências em bases de dados com um número elevado de faces a analisar. Finalmente, ao nível de EGM, foi precursor o trabalho realizado por Wiskott *et al.* [?].

Tabela 2.1: Categorização dos Métodos de Reconhecimento Facial em Imagens por Zhao *et al.* [ZCPR03]

Método	Trabalho Representativo
Holistic methods	
<i>Principal Component analysis (PCA)</i>	
Eigenfaces	Direct application of PCA [Craw and Cameron 1996; Kirby and Sirovich 1990; Turk and Pentland 1991]
Probabilistic eigenfaces	Two-class problem with prob. measure [Moghaddam and Pentland 1997]
Fisherfaces/subspace LDA	FLD on eigenspace [Belhumeur et al. 1997; Swets and Weng 1996b; Zhao et al. 1998]
SVM	Two-class problem based on SVM [Phillips 1998]
Evolution pursuit	Enhanced GA learning [Liu and Wechsler 2000a]
Feature lines	Point-to-line distance based [Li and Lu 1999]
ICA	ICA-based feature analysis [Bartlett et al. 1998]
<i>Other Representations</i>	
LDA/FLD	LDA/FLD on raw image [Etemad and Chellappa 1997]
PDBNN	Probabilistic decision based NN [Lin et al. 1997]
Feature-based methods	
Pure geometry methods	Earlier methods [Kanade 1973; Kelly 1970]; recent methods [Cox et al. 1996; Manjunath et al. 1992]
Dynamic link architecture	Graph matching methods [Okada et al. 1998; Wiskott et al. 1997]
Hidden Markov model	HMM methods [Nefian and Hayes 1998; Samaria 1994; Samaria and Young 1994]
Hybrid methods	
Convolution Neural Network	SOM learning based CNN methods [Lawrence et al. 1997]
Modular eigenfaces	Eigenfaces and eigenmodules [Pentland et al. 1994]
Hybrid LFA	Local feature method [Penev and Atick 1996]
Shape-normalized	Flexible appearance models [Lanitis et al. 1995]
Component-based	Face region and components [Huang et al. 2003]

Desde então, diversos investigadores estenderam estes três tipos de algoritmos. Em 2003, Zhao *et al.* [ZCPR03] efetuaram um estudo aprofundado das técnicas desenvolvidas até então classificando-as em três categorias. Métodos holísticos, caso estes usem toda a região da face como fonte de comparação no reconhecimento, métodos baseados nas características faciais, caso estes usem as características extraídas como informação essencial para a classificação da face e ainda como híbridos, caso seja utilizada uma mistura de ambos os métodos. Os resultados desta classificação podem ser visualizados na tabela 2.1.

Tal como referido na secção 2.1.2 deste relatório, o reconhecimento facial é uma tarefa complexa, composta por um conjunto de sub-problemas que enfrentam uma série de desafios para a sua conclusão. O primeiro desafio em qualquer sistema de reconhecimento facial é a deteção da face na imagem, a esse nível o trabalho realizado por Paul Viola e Michael Jones [VJ04] é considerado como um dos mais robustos métodos disponíveis na atualidade. Outras áreas onde a investigação atual se encontra focada têm a ver com a criação de métodos de reconhecimento facial robustos em relação à iluminação das imagens, assim como a pose dos indivíduos representados. [CSP10].

Ao nível da pose, o uso de *3D morphable face models* [BV03] tem revelado resultados positivos em posições não frontais. Estes algoritmos simulam o processo de formação de uma face em 3D, estimando a sua representação tridimensional a partir de imagens 2D fornecidas ao sistema com recurso a técnicas de computação gráfica. Esta estimativa é efetuada através do encaixe de um modelo 3D de faces na própria imagem fornecida, modelo esse apreendido através da digitalização 3D de um conjunto de rostos e respetiva textura. No entanto, em grande parte dos modelos desenvolvidos a seleção manual de um pequeno número de características faciais é requerida e o poder de computação necessário para a criação dos modelos é elevado. Extensões destes modelos foram também propostas para criar sistemas mais robustos em relação à variação na iluminação.

Outras abordagens que têm em vista a criação de sistemas de reconhecimento facial robustos em relação à iluminação centram os seus esforços na estimação do albedo para a normalização das imagens. O albedo representa a relação entre a quantidade de luz refletida por um ponto e quantidade de luz incidente, métodos recentes centram-se no desenvolvimento de filtros estocásticos não estacionários para a estimação de mapas de albedo das amostras faciais [BAC09]. Estes mapas podem ainda ser combinados com os modelos 3D faciais desenvolvidos para uma representação facial mais completa. Apesar de os resultados obtidos por estes métodos serem promissores quando comparados com os métodos tradicionais como o *Eigenfaces*, estes métodos ainda carecem de maior validação, uma vez que as suas avaliações foram feitas utilizando conjuntos de dados controlados e de reduzida dimensão [CSP10].

O investimento realizado na investigação de metodologias de reconhecimento facial automático tem permitido progressos notáveis na área. Do mesmo modo, a investigação em áreas relacionadas como a deteção de faces nas imagens, o envelhecimento facial ou o reconhecimento de expressões faciais em imagens tem permitido evoluir ainda os resultados obtidos. Devido à complexidade envolvida no processo de reconhecimento facial automático a evolução na área encontra-se assim intrinsecamente ligada à evolução verificada nas múltiplas áreas envolvidas, sendo que a investigação independente de cada uma é crítica para uma resolução eficaz da grande variabilidade de desafios que o reconhecimento facial automático se propõe ultrapassar.

### 2.3.1 Soluções Existentes

A investigação desenvolvida nos vários centros de investigação na década de 90, levou à passagem das soluções de reconhecimento facial automático de meros protótipos de laboratório, para soluções comerciais de elevado valor acrescentado, dessas soluções resultaram três categorias comerciais de produtos na área do reconhecimento facial:

- Soluções Completas. Estas soluções caracterizam-se por oferecer toda a infraestrutura necessária para efetuar o reconhecimento facial, quer a nível de software, quer a nível de hardware.
- Software. Soluções tradicionais de sistemas de reconhecimento facial, que consistem em software que permite efetuar o reconhecimento facial independentemente do hardware utilizado para a captura de informação. Dentro destas soluções existe uma grande variedade nos produtos oferecidos, desde soluções baseadas na captura de imagens 2D estáticas, imagens 3D e vídeo. Para além disso, as soluções oferecidas encontram disponíveis para serem utilizadas em diversas plataformas, como por exemplo dispositivos móveis ou através da Internet.
- SDK para reconhecimento facial. Permitem aos clientes, através de módulos independentes ou APIs, a criação de novas aplicações de reconhecimento facial.

Para além do reconhecimento facial automático, as soluções disponíveis possuem muitas vezes também associados outros métodos de biometria, como por exemplo reconhecimento de íris ou impressões digitais, assim como métodos mais tradicionais de autenticação como o uso de cartões de identificação ou passwords, de modo a aumentar a robustez e segurança dos sistemas.

Na tabela ??, encontram-se listadas algumas das soluções comerciais atualmente existentes de reconhecimento facial, assim como alguns exemplos de aplicações onde estas foram utilizadas. Esta tabela não pretende representar um levantamento exaustivo das soluções existentes mas apenas ilustrar a vasta gama de produtos disponíveis.

Para além das soluções comerciais de reconhecimento facial ao longo dos últimos anos surgiu também algumas soluções gratuitas de reconhecimento facial. Deste tipo de soluções são exemplo o software Picasa da Google, ou software iPhoto da Apple que permitem aos seus utilizadores a organização de álbuns multimédia com recurso ao reconhecimento facial. Por outro lado, redes sociais como Google Plus e o Facebook, lançaram também recentemente serviços que permitem aos seus utilizadores a deteção e identificação automática de outros utilizadores presentes nas imagens carregadas para os seus servidores.

## 2.4 Evolução Desempenho

A performance de um sistema de reconhecimento facial depende de uma grande variedade de fatores como a iluminação, posição da face na imagem, expressão facial e presença ou ausência de adereços (ex:óculos). Por outro lado, as condições de utilização do próprio sistema, tais como o número de utilizadores diferentes e a frequência de utilização, são também fatores a ter em conta ao avaliar a performance de um sistema. Uma avaliação apropriada dos sistemas de reconhecimento facial encontra-se então dependente da existência de uma base de dados de grandes dimensões e consequentemente um conjunto de casos de teste vasto, assim como, da existência de métodos apropriados para essa avaliação.

A criação do programa FERET em 1993, pelo NIST, permitiu assim colmatar uma lacuna existente na avaliação dos sistemas de reconhecimento facial existentes com a criação de uma base de dados de faces de grande dimensões assim como de um protocolo de avaliação dos respetivos sistemas de reconhecimento facial. Com base no protocolo desenvolvido foram levadas a cabo avaliações aos nos anos de 1994 e 1995. Um novo protocolo foi criado para a avaliação realizada em 1996. Posteriormente, o mesmo instituto realizou novas avaliações baseadas no protocolo FERET96 na edição de 2000 dos FRVT, a partir de 2002, um novo protocolo foi criado (protocolo FRVT 2002), com base no protocolo FERET96, mas com a particularidade de permitir a avaliação de sistemas de biometria em geral e não apenas sistemas de reconhecimento facial. Com base no novo protocolo criado foram então realizadas avaliações nos anos de 2002, 2006 e 2010. A decorrer encontram-se ainda as avaliações FRVT 2012.

As avaliações levadas a cabo durante os programas FERET e FRVT constituem desta forma uma boa base para a análise da evolução dos sistemas de reconhecimento facial automático ao longo dos últimos 20 anos, tal como pode ser visto na figura 2.2. De seguida encontram-se descritas os principais objetivos e conclusões retiradas dos programas FERET e FRVT:

### 2.4.0.1 FERET

Para além dos objetivos anteriormente citados de criação de uma base de dados de grandes dimensões e de criação de um protocolo que permitisse uma avaliação concreta e de base científica dos algoritmos existentes, o programa FERET tinha ainda como propósitos avaliar o estado da arte e identificar futuras áreas de desenvolvimento do reconhecimento facial [PRR00]. Tendo em conta estes objetivos, foi então avaliada a performance dos diferentes algoritmos em diferentes cenários, categorias de imagens e diferentes versões dos próprios algoritmos. A performance foi computada para situações de identificação e verificação de faces, estando os resultados da última avaliação efetuada, em 1996, descritos em Phillips *et al.* [PRR00] para os casos de identificação, e em Rizvi *et al.* [?] para os situações de verificação de faces.

A primeira avaliação do programa FERET, em 1994, consistia num três de conjuntos de testes, cada um com uma galeria e conjunto de provas diferentes. O primeiro conjunto, consistia numa prova de reconhecimento de faces de uma galeria de 316 faces. O segundo, tinha com objetivo determinar a capacidade dos algoritmos de rejeitar faces não presentes na galeria, designadas de falso-alarme. O terceiro, tinha como objetivo analisar os efeitos da pose na performance dos algoritmos [PRR00].

Em 1995, teve lugar a segunda avaliação, com o objetivo de avaliar os mesmo algoritmos em galerias de maior dimensão. A performance foi medida num único teste, com uma galeria de 817 indivíduos conhecidos, tendo sido dado um maior ênfase na avaliação de provas duplicadas. Um duplicado é uma imagem cuja imagem correspondente da galeria foi geralmente capturada em dias diferentes [PRR00].

Na última edição, em 1996, e já com um protocolo atualizado, foram testados 12 algoritmos de reconhecimento facial, e estudada a sua performance em quatro situações diferentes, usando para isso uma galeria com um total de 1196 indivíduos. Na primeira situação, a galeria e as

imagens de prova foram tiradas no mesmo dia e sobre as mesmas condições de iluminação. No segundo conjunto de provas, a galeria e as imagens de prova foram tiradas em dias diferentes. Para o terceiro teste, a galeria e as respetivas provas foram capturadas com pelo menos um ano de diferença. Finalmente, no quarto e último conjunto de testes, a galeria e as imagens de prova foram tiradas no mesmo dia, mas com diferentes condições de iluminação.

Várias conclusões foram retiradas das avaliações efetuadas. A primeira conclusão foi que, de facto, se verificava uma evolução nas técnicas de reconhecimento facial, uma vez que os resultados foram progressivamente melhores, mesmo quando comparadas versões diferentes do mesmo algoritmo. Essa conclusão é suportada pela análise dos resultados do caso geral de identificação, testado em todas as edições, onde as imagens da galeria e da prova foram capturadas no mesmo dia e com a mesma iluminação. Nesse teste em 1994, apenas 78% indivíduos foram identificados com sucesso numa galeria de 317 indivíduos, ao passo que em 1995 foram identificados já 93 % dos indivíduos numa galeria de 831 e nos testes da última edição (1996), 95% dos indivíduos foi identificado numa galeria de 1196 indivíduos. De igual forma, foi também verificada evolução na deteção de duplicados, mantendo-se no entanto as percentagens de identificação de duplicados mais reduzidas quando comparadas com o caso geral de identificação devido à maior complexidade do problema.

Outra conclusão retirada foi que a performance se encontra dependente da galeria e conjunto de imagens de prova utilizadas, uma vez que foi verificado que mudando a galeria, mas mantendo o mesmo tipo de teste a ser efetuado a percentagem de indivíduos identificados varia significativamente. Para a galeria e conjunto de prova capturadas no mesmo dia a percentagem varia entre os 80% e os 94%, enquanto que para galeria e conjunto de prova, capturadas em dias diferentes essa percentagem varia entre os 24% e os 69% porcento. Para além disso, não se existiu qualquer algoritmo que se revelasse superior em todas as quatro situações de teste realizadas em 1996, variando as performances, conforme a situação de teste.

Finalmente, conforme a tarefa a efetuar, identificação ou verificação de faces, ficou também demonstrado que não havia uma técnica que se superiorizasse em ambas as situações, comprovando que o reconhecimento facial é uma tarefa complexa que necessita de um estudo e investigação concretos tendo em conta a área de aplicação em vista.

### 2.4.0.2 FRVT

Durante a década de 90 registou-se um progresso assinalável nas diversas técnicas de reconhecimento facial, tendo para isso sido muito importante o papel desempenhado por iniciativas como o FERET. No final da década de 90, assistiu-se então a uma passagem dos sistemas de reconhecimento facial existentes de meros protótipos universitários, para aplicações comerciais de reconhecimento facial automático, tal como pode ser comprovado pela participação de 5 soluções comerciais na primeira edição dos *Facial Recognition Vendor Test*, no ano 2000. Esta avaliação, tinham como objetivo maioritário efetuar uma avaliação técnica das capacidades dos sistemas comerciais de reconhecimento facial disponíveis, de forma a compreender os pontos fortes e fracos

de cada sistema, obtendo também uma análise do estado da arte do reconhecimento facial à data [BBP01]. Estas avaliações foram ainda repetidas em 2002, 2006, 2010 e ainda em 2012.

A primeira edição dos FRVT, em 2000, utilizou o mesmo protocolo de avaliação da avaliação última avaliação FERET em 1996, no entanto, os testes efetuados eram significativamente mais exigentes do que a avaliação anterior, sendo utilizada também uma maior variedade de imagens. Os resultados da avaliação efetuada foram ser reportados para as seguintes oito categorias: compressão, distância, expressão, media, iluminação, pose, resolução e variação temporal. Ao nível das conclusões retiradas, podem ser destacados os seguintes tópicos [BBP01, CSP10, LJ11]:

- Compressão de imagens, utilizando JPEG, até 40:1 não reduziu a taxa de reconhecimento;
- Pose não afeta o reconhecimento de forma significativa até  $\pm 25^\circ$ , mas afeta significativamente quando são atingidos os  $\pm 40^\circ$ ;
- De forma equivalente ao registado no programa FERET, o uso de diferentes iluminações interiores não afeta significativamente os resultados obtidos, no entanto esses resultados são afetados caso haja uma mudança entre zonas interiores e exteriores;
- Registou-se uma melhoria significativa dos resultados obtidos em imagens tiradas com mais de 18 meses de distância, quando comparado com os resultados obtidos no FERET. (Aumento de 5% a 12% de faces identificadas conforme o conjunto de dados avaliado).

A segunda edição dos FRVT teve lugar em 2002, contou com a participação de soluções de reconhecimento facial de 10 empresas distintas. Esta edição consistiu em duas provas fundamentais, os *High Computational Intensity test (HCInt)* e os *Medium Computational Intensity test (MCInt)*. Os HCInt tinham como objetivo medir a performance dos sistemas existentes em 121 589 imagens frontais de 37 435 pessoas, representando um desafio computacional de elevado grau de exigência, como comprovado pelas 242 horas de computação necessárias para efetuar os testes. Os *Medium Computational Intensity test (MCInt)*, assemelhavam-se às medições efectuadas em avaliações anteriores, como o FERET e o FRVT 2000, tendo como objetivo medir a performance dos sistemas em imagens de diferentes categorias, nomeadamente não frontais, capturas no interior e exterior, entre outros. As conclusões retiradas desta avaliação foram as seguintes [CSP10, LJ11]:

- Melhores sistemas são insensíveis a variações de luz interior;
- Reconhecimento facial em imagens exteriores é ainda um desafio por resolver;
- Utilização de *3D morphable models* melhora a performance em situações não frontais;
- Características como sexo e idade podem afetar significativamente a performance, sendo que se registou maior facilidade em distinguir homens e pessoas com idade superior.
- O tempo de captura entre provas afeta significativamente a performance, registando-se uma diminuição linear da performance com o aumento do tempo.

## Reconhecimento Facial em Imagens

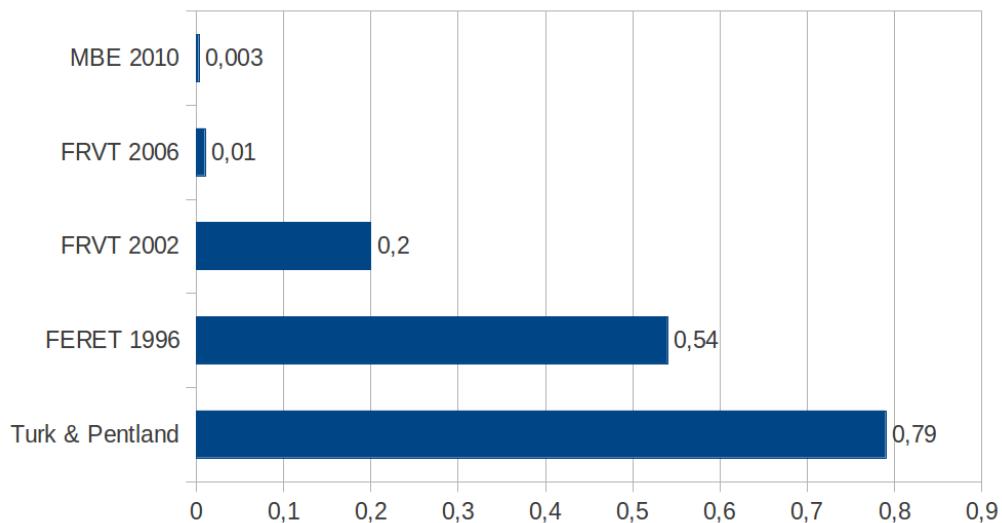


Figura 2.2: A redução da taxa de erro para os algoritmos estados da arte tal como documentado nas avaliações FERET, FRVT 2002, FRVT 2006 e MBE 2010 [PSO07, GQP10]

As últimas edições dos FRVT para as quais foram divulgados resultados foram as edições de 2006 [PSO07] e a de 2010. Na edição de 2010 a competição não assumiu o nome de FRVT 2010, mas sim de *Multiple Biometric Evaluation (MBE) 2010-still image track*, uma vez que além de avaliados sistemas de reconhecimento facial automático foram também avaliados sistemas de biometria com base na análise da íris. [GQP10]. Foi também realizada uma edição dos FRTV em 2012, mas a avaliação dos sistemas ainda se encontra a decorrer, não havendo por isso resultados divulgados.

Em 2006, foi analisada a performance de sistemas de reconhecimento facial, de 22 organizações e de 10 países diferentes, baseados em representações 3D e em imagens estáticas de faces. As avaliações consistiam em três testes distintos: comparação de imagens capturadas com luz interior; comparação de digitalizações de faces 3D com base na forma e textura; comparação de imagens capturadas em estúdio com imagens capturadas em corredores e átrios. Em 2010, para além do conjunto de dados analisados em 2006 e 2002, de modo a averiguar a evolução comparativa dos algoritmos, foi também analisado um segundo conjunto de imagens recolhidas por várias agências de segurança e compiladas pelo *Federal Bureau of Investigation (FBI)*. Este segundo conjunto de dados possuía mais de 1 milhão de faces fazendo da avaliação de 2010 a avaliação em maior escala feita aos sistemas de reconhecimento facial à data. Das provas de 2006 e 2010, as principais conclusões obtidas foram as seguintes [PSO07, GQP10, CSP10, LJ11]:

- Desde 1993 foram atingidas melhorias de 2 ordens de magnitude;
- Os resultados de 2010 e 2002 mostram uma melhoria de 1 ordem de magnitude nos sistemas de reconhecimento facial (Para um *false accept rate (FAR)* de 0.001, foi registado em 2002 um *false reject rate (FRR)* de 0.2, enquanto que em 2010 *FRR*=0.038);

## Reconhecimento Facial em Imagens

- Em condições de captura de informação controlada, a performance medida de sistemas de reconhecimento facial e sistemas baseados na análise da íris foi equivalente;
- A performance de sistemas baseados em imagens estáticas e sistemas baseados em representações 3D da face foi equivalente;
- Iluminação e resolução das imagens representam um papel fundamental no reconhecimento facial;

## Capítulo 3

# Abstração de Imagens

Os conteúdos foto-realistas, possuem muitas vezes mais informação do que a necessária para uma comunicação visual eficaz. A área do *Image Based Artistic Rendering (IB-AR)* tem como objetivo fazer a transformação de imagens e vídeos foto-realistas 2D em novas formas de representação visuais estilizadas artisticamente. A estilização de imagens e vídeos permite simplificar o conteúdo visual, assim como estrutura-lo de forma a dar ênfase e enfoque a determinadas características das imagens. Desta forma, é possível guiar a percepção obtida dos conteúdos visuais conforme o objetivo da mensagem a comunicar. Por outro lado, as técnicas de IB-AR permitem criar novas formas de representação gráfica de informação, dotando assim os artistas de novas ferramentas de comunicação e de criação de conteúdos visuais.

Atualmente a área da IB-AR diversificou-se numa atividade multidisciplinar que engloba áreas como a visão por computador, interação humano computador, computação gráfica e a modelação perceptual<sup>1</sup>. Os sistemas de IB-AR evoluíram também desde sistemas semi-automáticos de processamento de imagem no início dos anos 90, até sistemas totalmente automáticos de renderização. Adicionalmente o recurso a técnicas de visão por computador e o uso de filtros de abstração de imagens, permitiram melhorias notáveis ao nível da manipulação de grandes quantidades de informação, assim como da informação visual produzida. Uma evolução desta técnicas até atualidade pode ser vista na imagem 3.1.

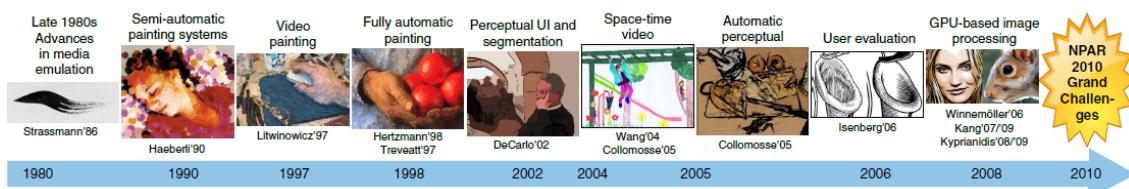


Figura 3.1: Cronologia do desenvolvimento da área da IB-AR [KCWI12].

As técnicas de IB-AR existentes podem ser divididas nas seguintes quatro categorias [KCWI12]:

<sup>1</sup>verificar *perceptual modeling*

- *Stroke Based Rendering*, as quais criam iterativamente camadas virtuais de traços imitando um pincel, cujas cores, orientação e escala são derivadas de processos automáticos ou semi-automáticos de processamento de imagem;
- *Region Based Techniques*, que usam a segmentação da imagem em diversas regiões para a sua posterior estilização;
- Renderização Baseada em Exemplo, que tentam imitar estilos artísticos específicos através de heurísticas apreendidas para a renderização automática de novas imagens nesses mesmos estilos;
- Filtros de Abstração de Imagens, os quais usam técnicas de processamento de imagem para, através de operações de filtragem local, efetuarem a renderização artística e abstração do conteúdo visual.

Apesar de os filtros de abstração de imagens possuírem uma menor diversidade numa perspectiva artística de renderização, a utilização de uma abordagem local por parte destes filtros torna-os vantajosos, quando comparada com as restantes técnicas de renderização artística de imagens. Esta abordagem permite uma fácil adaptação dos processos de abstração de imagens a tecnologias de processamento paralelo, presentes nos CPUs modernos, assim como uma implementação direta do processo de abstração de alguns dos filtros pelos próprios GPUs. Nesse sentido, e devido ao grande número de imagens utilizadas no processo de reconhecimento facial automático a utilização de filtros de abstração de imagens torna-se mais vantajosa no âmbito desta dissertação.

### 3.1 Filtros de Abstração de Imagens

Os filtros de abstração de imagens permitem simplificar o conteúdo visual, destacando apenas a informação essencial contida nas imagens. Existem vários tipos de filtros de abstração, de seguida encontram-se descritas cinco categorias de filtros de abstração de imagem tipicamente utilizados para a IB-AR.

#### 3.1.1 Filtro Bilateral

Uma abordagem baseada num filtro bilateral foi proposta em primeira mão por Tomasi e Manduchi em 1998 [TM98] e encontra-se representada na imagem 3.2b. Já em 2006, Winnemöller propôs uma melhoria dos resultados obtidos através da associação de um filtro de diferença de gaussianas na aplicação do filtro bilateral [WOG06]. Trabalhos posteriores com vista a melhorar, os contornos produzidos pelas imagens abstraídas foram também realizados com sucesso através da introdução de um filtro de diferença de gaussianas baseado em fluxo [?].

O filtro bilateral caracteriza-se por suavizar as zonas com menor contraste da imagem enquanto preserva os limites de maior contraste. No entanto, os efeitos deste filtro são poucos visíveis quando toda a imagem possui um elevado contraste. Por outro lado, em imagens pouco

## Abstração de Imagens

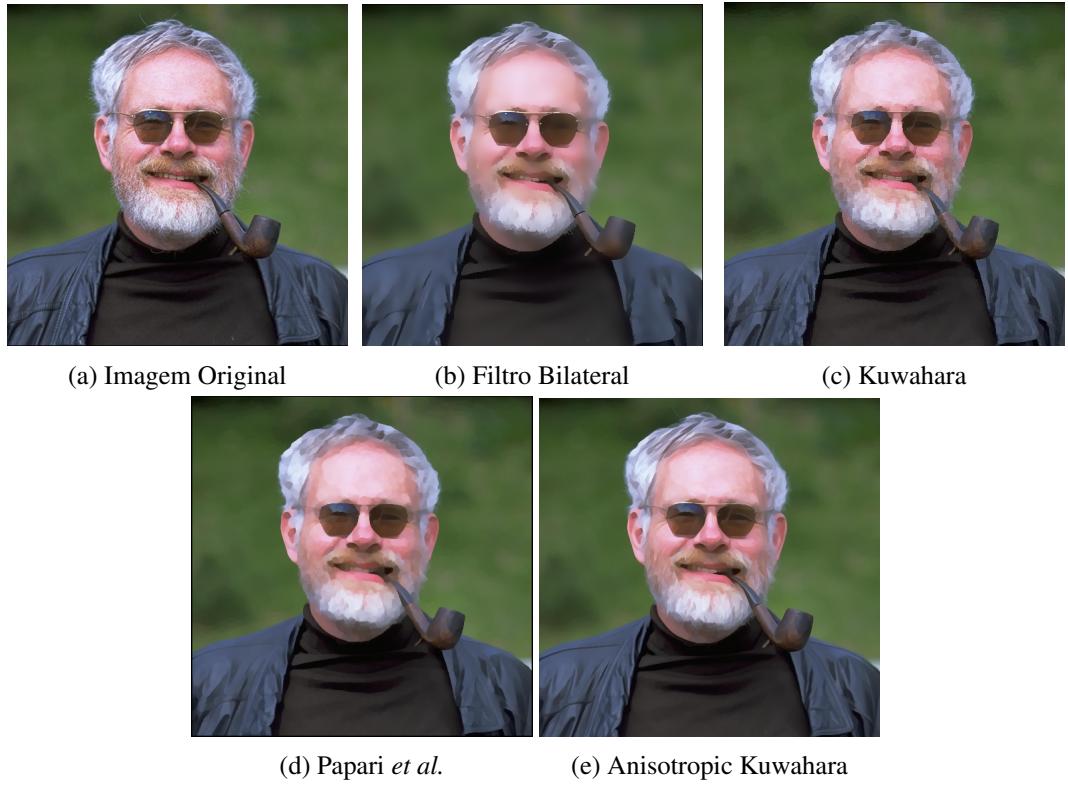


Figura 3.2: Diferentes filtros de abstração aplicados sobre a mesma imagem [KKD09].

contrastadas a informação visual removida é por vezes demasiada, causando um efeito de imagem desfocada na imagem abstraída.

### 3.1.2 Kuwahara

O filtro Kuwahara (fig. 3.2c) e as suas variantes, representam uma classe de filtros onde se tem verificado uma pesquisa ativa ao longo das últimas décadas [KKD09].

A ideia geral do filtro do tipo Kuwahara consiste em efetuar um achatamento da imagem, através da divisão da área a filtrar em quatro sub-regiões retangulares de menor dimensão, e a posterior substituição da informação contida numa dessas regiões pela média da sub-região com menor variância. Desta forma, é removido o detalhe em zonas com elevado contraste mas são preservadas as limites das formas presentes na imagem. No entanto, na presença de ruído este filtro revela-se instável, criando artefactos na imagem. Por último, esta abordagem revela-se também agressiva na presença de anisotropia<sup>2</sup>, ou seja quando a informação direcional presente na imagem é forte (por exemplo: pelo, cabelos), não preservando essa informação na imagem abstraída.

Existem várias variantes ao filtro Kuwahara, sendo que a grande diferença entre a maior parte das variantes consiste na forma como são definidas as sub-regiões a filtrar. Papari *et al.* [PPC07], apresentaram, com bons resultados, uma variante deste filtro em que a divisão das sub-regiões é efetuada em regiões circulares, que por sua vez se encontram divididas em oito regiões de igual

---

<sup>2</sup>Anisotropia - qualidade de certos materiais cujas propriedades são diferentes consoante as direções.



Figura 3.3: *Diffusion Shock Filter* [KL08].

tamanho. Apesar das melhorias registadas na abordagem de Papari *et al.*, visíveis na figura 3.2d, esta abordagem continua a não preservar a informação direcional contida nas imagens resultando assim no aparecimento de alguns artefactos, ainda que mais pequenos e com maior definição do que na solução proposta pelo o filtro original.

Para ultrapassar as limitações do filtro Kuwahara original na presença de anisotropia, foi proposta por Kyriyanidis *et al.* [KKD09], uma nova abordagem com o nome filtro Kuwahara Anisotrópico (3.2e). Nesta solução, é feita uma adaptação da forma, orientação e escala do método de seleção das sub-regiões à estrutura local da imagem, resultando numa forma que varia entre um círculo e um elipse alongada conforme a própria estrutura local da imagem. Assim, para além de preservada, a informação direcional é ainda enfatizada, resultando em imagens tipo pintura, com limites bem definidos e sem grandes artefactos visíveis.

### 3.1.3 *Diffusion Shock Filter*

Kang e Lee [KL08], foram os primeiros a utilizar *shock filtering* para IB-AR através do uso de *Mean Curvature Flow (MCF)*, resultando em imagens onde para além de suavizadas as variações irrelevantes das cores, são simplificados os limites das formas presentes nas mesmas, como é possível visualizar na figura 3.3. O resultado da aplicação de *shock filtered MFC*, é, no entanto, tipicamente muito agressivo, resultando em imagens desfocadas, onde apenas é parcialmente preservada a informação direcional contida na imagem. Esta técnica revela-se também, pouco indicada para a abstração de vídeo, uma vez que é pouco sensível a pequenas alterações nas imagens.

### 3.1.4 Filtragem Morfológica

Os filtros desta categoria, utilizam os princípios da morfologia matemática<sup>3</sup> para reduzir o detalhe existentes nas zonas coloridas, criando um efeito de uniformização da cor.

### 3.1.5 Técnicas *Gradient Domain*

O gradiente de uma imagem representa de forma direcional as mudanças de intensidade de uma imagem independente das suas cores originais, permitindo assim uma maior flexibilidade na manipulação das imagens. As técnicas baseadas no uso de gradientes, consistem na construção de um campo de gradiente que representa a imagem, o qual é posteriormente utilizado na reconstrução da imagem filtrada.

A grande desvantagem das técnicas baseadas no uso de gradientes é que os resultados obtidos em zonas com contraste são por vezes problemáticos e requerem correção. Para além disso, a computação deste tipo de filtragem é computacionalmente pesada e por isso não aplicável para situações de utilização em tempo real.

## 3.2 Trabalho Preliminar

Os filtros de abstração são normalmente utilizados por artistas para comunicar a mensagem visual mais eficazmente. No âmbito desta dissertação pretendemos analisar o impacto destes filtros não ao nível da comunicação visual efetuada, mas ao nível do impacto que estes podem causar no processo de reconhecimento facial. Esta investigação é motivada por resultados anteriores obtidos com o uso dos mesmos filtros na ilustração automática de texto.

A ilustração de texto é uma tarefa enquadrada na área da recuperação de informação multimédia e que consiste na pesquisa de imagens adequadas para a ilustração de fragmentos de texto, tais como notícias ou livros. Coelho e Ribeiro [CR12], propõem o uso da abstração de imagens para a ilustração automática de texto, através de um sistema de recuperação de informação, o qual dada uma entrada textual inicial, seleciona automaticamente uma pequena lista de imagens relacionadas a partir da análise coleção de imagens previamente processada.

O sistema desenvolvido utiliza a abstração de imagens para melhorar a informação capturada pelos descritores de informação visual utilizados e consequentemente melhorar os resultados obtidos a partir da análise baseada em conteúdos efetuada à coleção de dados.

Os resultados obtidos demonstraram que o uso de abstração tem a potencialidade de melhorar a informação recuperada, assim como reduzir as necessidades de processamento e armazenamento da coleção de dados utilizada.

---

<sup>3</sup>Mathematical morphology (MM) is a theory and technique for the analysis and processing of geometrical structures, based on set theory, lattice theory, topology, and random functions.

## Capítulo 4

# Perspectiva de Solução

O reconhecimento facial em imagens sofreu uma evolução notável nos últimos 20 anos, tal como documentado no capítulo 2 deste relatório.

Em cenários cooperativos com condições de captura de imagens controladas, nomeadamente ao nível da pose, iluminação e expressões faciais, considera-se mesmo que o problema de verificação 1:1 se encontra praticamente resolvido, uma vez que as taxas de reconhecimento atingidas são satisfatórias para a grande maioria das aplicações [LJ11]. Existem também várias aplicações em situações reais com um bom nível de satisfação por parte dos seus utilizadores, como é o caso do sistema de fronteira automático dos aeroportos portugueses (ver 2.2.2), ou o controlo de entradas nas cerimónias inaugurais dos jogos olímpicos de Pequim (ver ). Em condições específicas e favoráveis, é então possível considerar que os sistemas de reconhecimento facial automático atuais conseguem mesmo ultrapassar a capacidade de reconhecimento humana, uma vez que conseguem identificar com precisão um maior número de faces do que aquelas que um humano consegue.

Contudo, o problema de reconhecimento facial automático ainda se encontra longe de ser um problema totalmente resolvido. Em cenários onde é registada uma grande variação ao nível da pose, iluminação ou outros fatores identificados na secção 2.1.2 deste relatório, a identificação das entidades capturadas é ainda uma tarefa desafiante [LJ11]. Para além disso, a performance e satisfação obtida por parte dos utilizadores dos sistemas atuais demonstra uma grande variação tendo em conta as situações onde estes sistemas são utilizados.

Por outro lado, a crescente ubiquidade tecnológica e poder computacional presente nos diversos dispositivos utilizados no nosso dia a dia, aumenta o leque de aplicações possíveis do reconhecimento facial automático, apresentando novos desafios às soluções de atualmente existentes.(ver )

Assim sendo, torna-se pertinente a continuação da investigação na área do reconhecimento facial em imagens. Para além disso, o elevado valor comercial das soluções existentes e consequente falta soluções abertas permite alcançar uma boa visibilidade dos resultados obtidos.

Ao nível da abstração de imagens, estudos efetuados demonstraram que aplicação destes filtros na recuperação de informação multimédia, nomeadamente no âmbito da ilustração automática

de texto têm a potencialidade de melhorar a informação retornada, assim como reduzir significativamente as necessidades de processamento e armazenamento das imagens [?].

Tendo em conta a importância e a necessidade de investigação na área do reconhecimento facial automático, e uma vez que não existem estudos relativos à utilização de filtros de abstração no processo de reconhecimento facial, torna-se pertinente o estudo do seu impacto, no âmbito desta dissertação. Desta forma, a hipótese levantada no âmbito desta dissertação é então que o uso de abstração em imagens que vão ser alvo de reconhecimento facial pode melhorar o processo de reconhecimento facial automático em imagens.

## 4.1 Objetivos

O principal objetivo deste projeto visa o estudo do impacto de filtros de abstração visual de informação no processo de reconhecimento facial automático em imagens. Para isso, destacam-se o seguinte conjunto de objetivos parciais:

1. Desenvolvimento de um sistema de reconhecimento facial de personalidades;
2. Integração da abstração de imagens no sistema desenvolvido;
3. Avaliação dos resultados da abstração de imagens no processo de reconhecimento facial;

Da avaliação efetuada, para além da publicação desta dissertação, espera-se ainda a publicação de um artigo científico onde sejam descritos os resultados obtidos.

## 4.2 Implementação

Uma vez que não se pretende no âmbito desta dissertação efetuar investigação ao nível dos algoritmos de reconhecimento facial, mas sim ao nível do impacto do uso de imagens abstraídas nesses sistemas, a implementação desses algoritmos terá por base a utilização de soluções abertas de reconhecimento facial, nomeadamente a biblioteca *Open CV (Open Source Computer Vision)*.

Ao nível dos filtros de abstração o estudo será efetuado utilizando o filtro *Anisotropic Kuwahara*. Este filtro foi utilizado anteriormente, e com resultados positivos, em abstração de imagens para a recuperação de informação multimédia, pelo que se considera adequada a sua utilização no âmbito deste projeto.

Por último, ao nível das coleções de dados a analisar, temos em vista a utilização de duas coleções a primeira é uma coleção standard, chamada *Labeled Faces in the Wild* e que permite obter resultados comparativos com avaliações efetuadas anteriormente a sistemas de reconhecimento facial existentes. A segunda é a coleção de imagens Sapo Fama, qual agrupa uma base de dados de imagens de personalidades famosas nacionais e internacionais, representando uma possível área de aplicação do sistema desenvolvido.

#### 4.2.1 Sistema de Reconhecimento Facial Base - OpenCV Face Recognizer

O *OpenCV* é uma biblioteca de código aberto nas áreas de visão por computador e *machine learning*, onde se encontra disponível a implementação de mais de 2500 algoritmos. Esta biblioteca possui uma comunidade de mais de 47 mil pessoas, já registou mais de 5 milhões de downloads e é utilizada globalmente por empresas como a Google, Yahoo, Microsoft, Intel, IBM, Sony, Honda, Toyota [its].

Ao nível do reconhecimento facial, esta biblioteca disponibiliza um módulo denominado *Face Recognizer*, onde se encontram implementados os algoritmos *Eigenfaces*, *Fisherfaces* e *Local Binary Patterns Histograms*.

Tendo em conta a ampla utilização da biblioteca OpenCV e a sua constante atualização pela comunidade, assim como as facilidades providenciadas pelo módulo de reconhecimento facial, esta foi considerada a biblioteca ideal para utilizar como base de implementação do sistema de reconhecimento facial a criar. De seguida encontram-se brevemente descritos os três algoritmos implementados no módulo *Face Recognizer*:

##### 4.2.1.1 Eigenfaces

O método *Eigenfaces* foi introduzido por [BHK97] e tira partido da Análise dos Componentes Principais (ACP) para efetuar o reconhecimento facial automático.

A análise de componentes principais tem como objetivo determinar as relações existentes entre diferentes conjuntos de dados, nomeadamente ao nível das suas diferenças e semelhanças, tirando partido da redundância existente para criar uma representação reduzida dos dados sem que a perda de informação ocorrida seja significativa. As imagens faciais possuem uma grande redundância natural, o algoritmo *Eigenfaces*, através da análise dos componentes principais dessas imagens, efetua uma projeção das imagens faciais num sub-espacó onde se evidenciam apenas as variações entre as diversas caras conhecidas pelo sistema.

A redução do espaço de representação revela-se fulcral no problema de reconhecimento facial em imagens, devido à grande dimensionalidade exigida para a representação de uma face. Considerando, por exemplo, uma dada imagem de  $n \times m$  pixels de tons cíntenos, essa imagem poderia ser traduzida por um espaço vetorial de  $m = n \times m$  pixels, assim sendo, uma imagem de apenas  $256 \times 256$  pixels necessaria de um total de 65536 pixels para ser representada.

O processo de reconhecimento facial com recurso ao algoritmo *Eigenfaces* consiste então nos seguintes passos:

1. Aquisição de um conjunto de dados iniciais (conjunto de treino);
2. Projeção dos dados obtidos num sub-espacó de faces através da ACP;
3. Aquisição de uma imagem a reconhecer;
4. Projeção da face a reconhecer no sub-espacó do conjunto de treino, calculando as distâncias obtidas para cada face conhecida pelo sistema;

5. Determinar qual a face do conjunto de treino com menor distância à face a reconhecer;
6. Caso a distância para a face obtida seja menor do que um limite operacional estabelecido, a imagem é reconhecida como sendo essa pessoa, caso contrário, a face é identificada como sendo uma pessoa desconhecida pelo sistema.

Este método efetua assim uma abordagem holística ao problema de reconhecimento facial em imagens, uma vez que tem em consideração a representação facial como um todo, não fazendo a distinção entre pontos específicos da face como olhos, orelhas ou nariz para efetuar o reconhecimento facial. Uma vantagem deste tipo de representação é a reduzida sensibilidade ao ruído presente nas imagens [ZCPR03].

#### 4.2.1.2 Fisherfaces

A análise dos componentes principais visa determinar o sub-espaço onde se verifica uma maior variação entre um conjunto de imagens, no entanto, a variação na representação facial de uma pessoa encontra-se muitas vezes relacionada com mudanças de expressões faciais ou iluminação dos indivíduos. Assim sendo, o sub-espaço criado pelo algoritmo *Eigenfaces* não traduz muitas vezes as apenas as diferenças de identidade entre os diversos indivíduos, mas também as diferenças verificadas ao nível das condições de captura das imagens, mesmo considerando o mesmo indivíduo devido à sua tolerância reduzida em termos de representação das variações intra-indivíduo. O algoritmo *Fisherfaces* tenta resolver este problema, através da aplicação de um passo de *Linear Discriminant Analysis (LDA)* após a análise dos componentes principais, de forma a determinar mais corretamente as variações intra-classe existentes no conjunto de imagens a avaliar.

A LDA tenta maximizar as diferenças existentes entre diferentes indivíduos(inter-classe) e minimizar as variações entre imagens da mesma pessoa(intra-classe) de modo a obter uma representação mais robusta em termos de variação ao nível da iluminação.

#### 4.2.1.3 Local Binary Patterns Histograms

lorem ipsum

### 4.2.2 Filtros de Abstração - Filtro Anisotropic Kuwahara

Ao nível dos filtros de abstração o estudo será efetuado utilizado o filtro Anisotropic Kuwahara (FAK).

Este filtro consiste numa generalização do filtro Kuwahara que remove alguns artefactos originados na aplicação do filtro original através da adaptação da forma, escala e orientação do filtro à estrutura local das características da imagem [KKD09]. Desta forma é produzido um efeito de abstração tipo pintura, onde é removida informação não essencial em zonas de elevado contraste, enquanto são preservados os limites representados nas zonas de menor contraste, tal como demonstrado na figura 4.1. As imagens ficam assim com a clareza de uma ilustração, mas preservam



Figura 4.1: Comparação entre imagem não abstraída (esquerda) e imagem abstraída (direita).

a informação direcional tal como nas pinturas a óleo clássicas. Por outro lado, este filtro tira partido da placa gráfica para a realização da abstração das imagens, tornando-se assim particularmente indicado para o processamento de um elevado número de fotografias.

O filtro em questão foi também utilizado anteriormente, e com resultados positivos, em abstração de imagens para a recuperação de informação multimédia.

Tendo em conta os fatores apresentados consideramos que este filtro é o mais adequado para a utilização no âmbito deste projeto.

### 4.2.3 Coleções de dados

Ao nível das coleções de dados a analisar serão utilizadas duas coleções de dados Sapo Fama e *Labeled Faces in the Wild*.

#### 4.2.3.1 *Labeled Faces in the Wild*

A coleção *Labeled Faces in the Wild* (*LFW*) é uma base de dados fotográfica desenhada especificamente para o estudo do problema de reconhecimento facial, particularmente em situações onde as condições de captura das imagens não possuem restrições. Esta coleção possuí 13233 imagens, de 5749 pessoas diferentes, sendo que dessas pessoas 1680 possuem mais do que uma imagem na galeria. O uso desta galeria no âmbito desta dissertação permitirá efetuar uma análise dos resultados obtidos num conjunto de dados standard e já previamente analisado por outros investigadores, assim como agilizar os testes dos sistemas desenvolvidos uma vez que a coleção já se encontra preparada especificamente para o estudo do desempenho de sistemas de reconhecimento facial automático. Nas figuras 4.3, 4.4 e 4.5 encontram-se representados alguns exemplos das imagens contidas nesta coleção e respetivas anotações.



Figura 4.2: Exemplos de imagens da coleção Sapo Fama.

#### 4.2.3.2 Sapo Fama

A coleção Sapo Fama foi disponibilizada no âmbito da integração desta dissertação no laboratório de investigação da Sapo da Universidade do Porto. Esta galeria multimédia agrupa uma coleção de imagens utilizadas na plataforma online de notícias sobre personalidades públicas da Sapo, capturadas num grande número de eventos com condições de iluminação, pose e expressão facial muito variáveis. Para além disso, a cada imagem encontra-se ainda associada uma descrição textual. As imagens contidas nesta coleção representam uma oportunidade única de avaliação do sistema desenvolvido numa biblioteca de imagens real e com elevado valor comercial. Na figura 4.2 encontram-se representados alguns exemplos das imagens contidas nesta coleção.

Para a utilização deste conjunto de dados será ainda necessário efetuar a preparação do mesmo nomeadamente através da extração e segmentação das diversas entidades presentes em cada imagem, assim como da normalização das imagens em termos de dimensão.

### 4.3 Avaliação Desempenho

A avaliação da desempenho dos sistemas de reconhecimento facial é normalmente efetuada em duas situações específicas: identificação e verificação de faces. No âmbito da avaliação desses sistemas designam-se de *amostras biometricas* as capturas de características de uma pessoa que permitem efetuar o seu reconhecimento. Dependendo do sistema essas amostras podem ser apenas uma imagem, um conjunto de imagens ou até um vídeo. Designa-se ainda de *prova* a amostra biométrica apresentada ao sistema para ser reconhecida.

Para uma avaliação precisa dos sistemas automáticos de reconhecimento facial são necessários três conjuntos de imagens: O primeiro, conjunto  $\mathcal{G}$ , designado de galeria, contem amostras biométricas das pessoas já conhecidas pelo sistema. Um exemplo de imagens da galeria pode ser visto na figura ??



Figura 4.3: Galeria com faces conhecidas pelo sistema e respectivas anotações.

O segundo grupo de imagens é o conjunto  $\mathcal{P}_g$ , engloba o conjunto de provas de pessoas conhecidas pelo sistema, mas que são diferentes das amostras biométricas presentes na galeria. Um exemplo de imagens deste conjunto pode ser visto na figura 4.4, onde se encontram representadas imagens da cantora Norah Jones, mas que são diferentes das imagens presentes no conjunto  $\mathcal{G}$ .

O último conjunto é designado de  $\mathcal{P}_N$ , e corresponde ao conjunto de provas de pessoas não presentes na galeria. Essas imagens são também muitas vezes designadas de impostores, uma vez que servem para testar as situações em que o sistema identifica erradamente pessoas desconhecidas pelo sistema como pessoas presentes na galeria. Um exemplo deste conjunto de imagens encontra-se representado na figura 4.5

Quando uma prova  $p_j$  é apresentada ao sistema, essa prova é então comparada a cada amostra biométrica da galeria  $g_i$ , resultando dessa comparação o respetivo índice de similaridade (*similarity score*),  $s_{ij}$ . Este índice é designado de *match score*, caso  $g_i$  e  $p_j$  sejam amostras da mesma pessoa, caso não o sejam é designado de *nonmatch score*.

A função  $id()$  retorna a identidade de uma amostra biométrica, em que  $id(p_j) = id(g^*)$ , sendo  $g^*$  a única correspondência de  $p_j$  na galeria e  $s_{*j}$  o respetivo índice de similaridade.

De seguida encontra-se descrita brevemente a forma como é efetuada a avaliação da desempenho dos sistemas de reconhecimento facial para os casos de identificação e verificação.



Norah Jones

Norah Jones

Figura 4.4: Imagens de pessoas conhecidas pelo sistema, mas não presentes na galeria.

### 4.3.1 Identificação

Tal como referido no capítulo 2 deste relatório, em sistemas de reconhecimento facial automático o problema de identificação consiste na determinação da identidade de uma amostra biométrica (prova) fornecida ao sistema.

Na identificação de uma dada prova são calculados os índices de similaridade para todas as amostras na galeria, e ordenados os seus resultados. Uma prova  $p_j$  tem ranking  $n$  se o seu índice de similaridade for o enésimo maior índice de similaridade. Caso haja empates nos índices de similaridade é necessário resolver esses empates para determinar o ranking de uma prova, havendo para isso três métodos: otimista, pessimista e média. No método otimista, uma prova é associada ao ranking mais alto possível, obtido pelo número de índices estritamente maiores do que  $s_{*j}$  mais um ( $| > s_{*j} | + 1$ ). Na abordagem pessimista, um prova é associada ao ranking mais baixo possível, sendo esse devolvido pelo número de índices maiores ou iguais a  $s_{*j}$  mais um ( $| \geq s_{*j} | + 1$ ). Na média, tal como o seu nome indica é efetuada uma média entre os valores obtidos pelos métodos pessimista e otimista, sendo esta a abordagem mais utilizada.

O desempenho da identificação em sistemas de reconhecimento facial pode ser avaliada em taxa de deteção e identificação e taxa de falso-alarme:

**Taxa de deteção e identificação ( $T_{DI}$ ):** Percentagem de provas do conjunto  $\mathcal{P}_g$  que são corretamente identificadas. Uma prova é corretamente identificada caso o índice de similaridade para o seu *match score* seja maior do que um limite operacional pré-definido  $\tau$ .

$$T_{DI}(\tau, n) = \frac{|\{p_j : p_j \in \mathcal{P}_g, rank(p_j) \leq n, s_{*j} \geq \tau\}|}{|\mathcal{P}_g|} \quad (4.1)$$

Em que no caso de se pretender apenas determinar a identificação da pessoa é devolvido apenas o *top match*, ou seja o resultado com ranking  $n = 1$ . Caso contrário, podem ser devolvidos todos os primeiros  $n$  resultados com índice de similaridade superior ao limite operacional estabelecido.



Nicolas Kiefer

Tiger Woods

Figura 4.5: Imagens de pessoas não conhecidas pelo sistema.

**Taxa de falso-alarme ( $T_{FA}$ ):** Percentagem de impostores identificados erradamente, ou seja, percentagem de pessoas do conjunto de provas  $\mathcal{P}_{\mathcal{N}}$ , identificadas como alguém presente na galeria  $\mathcal{G}$ .

$$T_{FA}(\tau) = \frac{|\{p_j : p_j \in \mathcal{P}_{\mathcal{N}}, s_{ij} \geq \tau\}|}{|\mathcal{P}_{\mathcal{N}}|} \quad (4.2)$$

Um sistema de reconhecimento facial ideal teria uma taxa de deteção e identificação de 1.0 e uma taxa de falso-alarme de 0.0.

### 4.3.2 Verificação

Em situações de verificação da identidade, para além de uma amostra biométrica (*prova*) é também fornecida ao sistema a identidade da pessoa que se pretende identificar. Cabe ao sistema determinar se a identidade e amostra fornecidas pertencem à mesma pessoa. O sistema compara então a prova à amostra biométrica guardada na galeria, da pessoa a quem corresponde a identidade fornecida. Da comparação efetuada resulta o índice de similaridade entre a prova fornecida e os dados já armazenados no sistema.

Em termos formais, a verificação pode ser avaliada pela taxa de verificação e taxa de falso-alarme, as quais podem ser descritas da seguinte forma:

**Taxa verificação ( $T_V$ ):** Percentagem de provas do conjunto  $\mathcal{P}_{\mathcal{G}}$  que são corretamente verificadas. O sistema aceita como verdadeira a identidade fornecida caso o índice de similaridade calculado seja superior a um limite operacional fornecido.

$$T_V(\tau, n) = \frac{|\{p_j : p_j \in \mathcal{P}_{\mathcal{G}}, s_{ij} \geq \tau, id(g_i) = id(p_j)\}|}{|\mathcal{P}_{\mathcal{G}}|} \quad (4.3)$$

**Taxa de falsa-aceitação ( $T_{FA}$ ):** Percentagem de impostores aceites erradamente, ou seja, percentagem de pessoas do conjunto de provas  $\mathcal{P}_{\mathcal{N}}$ , verificadas como alguém presente na galeria  $\mathcal{G}$ .

$$T_{FA}(\tau) = \frac{|\{p_j : p_j \in \mathcal{P}_{\mathcal{N}}, s_{ij} \geq \tau\}|}{|\mathcal{P}_{\mathcal{N}}||\mathcal{G}|} \quad (4.4)$$

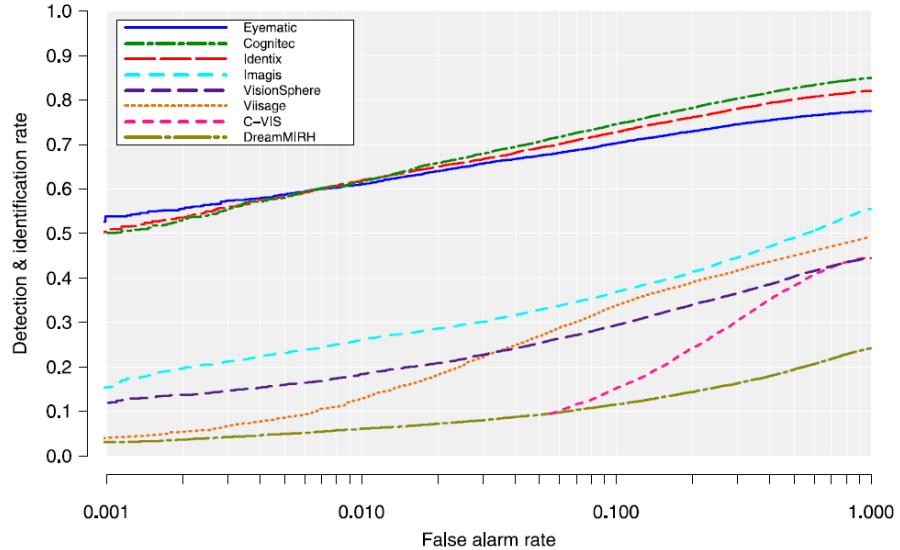


Figura 4.6: ROC.

A modificação do valor de  $\tau$  tem influência direta nas percentagens de deteção e/ou verificação e de falsos-alarmes calculadas. Ao aumentar o limite operacional, ambas as taxas diminuem, não sendo por isso possível maximizar ambos os valores, uma vez que existe sempre um compromisso entre o aumento da taxa de deteção e identificação e o aumento do número de falsos-alarmes. Este compromisso é tradicionalmente representado num gráfico do tipo *receiver operating characteristic* (ROC). Um exemplo de um ROC encontra-se representado na figura ??, com um ranking constante de 1.

Outra forma de representar a performance de sistemas de reconhecimento facial é em função do ranking das provas presentes na galeria e encontra-se representada na figura ???. Nesta figura, o eixo vertical representa a taxa de deteção e reconhecimento e o eixo horizontal o ranking numa escala logarítmica. Cada curva representa a performance do mesmo sistema com uma diferente taxa de falso-alarme.

## 4.4 Resultados Esperados

Tendo em conta os resultados anteriormente observados no recurso à utilização de filtros de abstração para a ilustração automática de texto e o estado da arte da área de reconhecimento facial automático em imagens, esperamos, no contexto desta dissertação, determinar o impacto do uso da abstração de imagens no processo de reconhecimento facial automático, nomeadamente nos seguintes pontos:

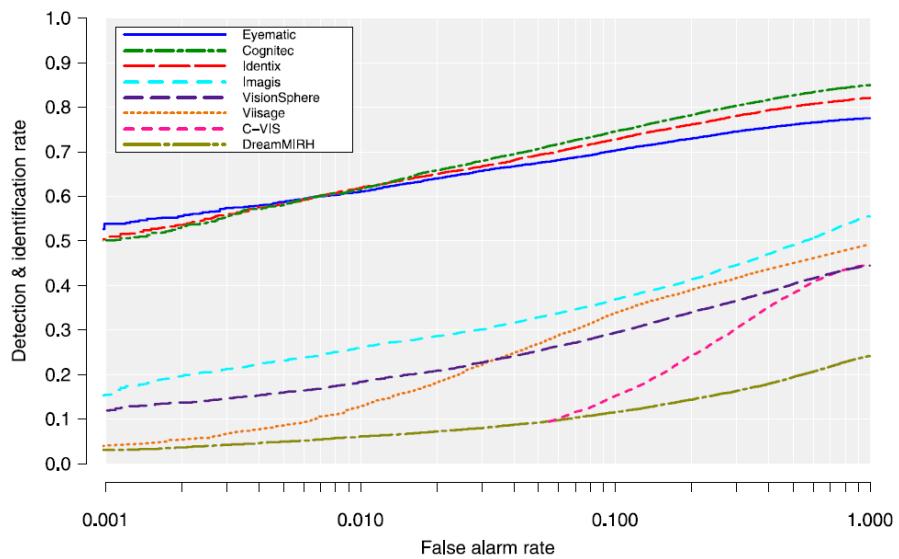


Figura 4.7: ROC2.

- Eficácia do Reconhecimento;
- Necessidades de Processamento das imagens;
- Necessidades de Armazenamento das imagens.

Em cada um dos pontos identificados anteriormente pretendemos determinar, qual o impacto ao nível na performance (melhoria ou perda de eficácia) com o uso de abstração de imagens, e qual o compromisso existente entre a eficácia do reconhecimento e as necessidades de processamento e armazenamento das imagens.

## 4.5 Plano de Trabalho

//Novo gant só com 2º semestre?

## Perspectiva de Solução

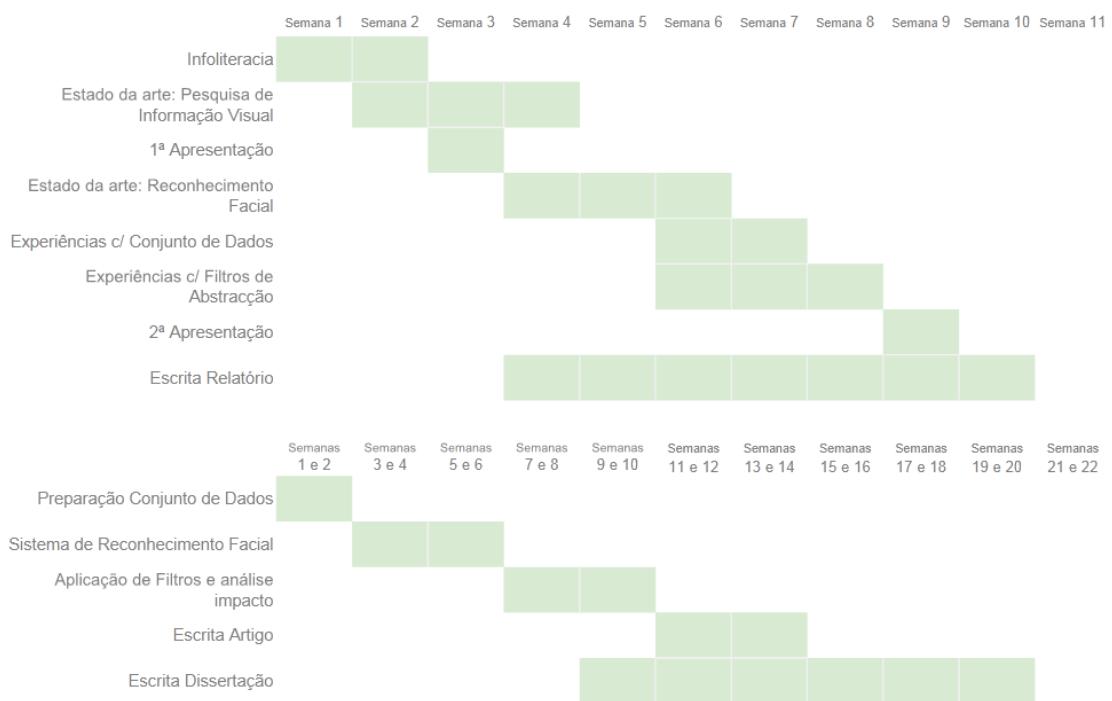


Figura 4.8: Plano de Trabalho 1º e 2º Semestres

## **Capítulo 5**

### **Conclusão**

TO-DO

# Referências

- [Age] Agencia para a Administracao Administrativa. Estatísticas Cartão Cidadão. URL: [http://www.cartao-decidadao.pt/index.php?option=com\\_content&task=view&id=295&Itemid=114&lang=pt](http://www.cartao-decidadao.pt/index.php?option=com_content&task=view&id=295&Itemid=114&lang=pt).
- [BAC09] Soma Biswas, Gaurav Aggarwal e Rama Chellappa. Robust estimation of albedo for illumination-invariant matching and shape recovery. *IEEE transactions on pattern analysis and machine intelligence*, 31(5):884–99, May 2009. URL: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=4531748](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4531748), doi:10.1109/TPAMI.2008.135.
- [BBP01] Duane M. Blackburn, Mike Bone e P. J. Phillips. Face Recognition Vendor Test 2000: Evaluation Report. page 71, February 2001. URL: <http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA415962>.
- [BHK97] P. N. Belhumeur, J. P. Hespanha e D. J. Kriegman. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):711–720, July 1997. URL: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=598228](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=598228), doi:10.1109/34.598228.
- [BT54] Jerome S. Bruner e Renato Tagiuri. THE PERCEPTION OF PEOPLE. page 61, September 1954. URL: <http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=AD0024982>.
- [BV03] V. Blanz e T. Vetter. Face recognition based on fitting a 3D morphable model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(9):1063–1074, September 2003. URL: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=1227983>, doi:10.1109/TPAMI.2003.1227983.
- [Chi] Chinese Academy Of Sciences. Facial recognition technology safeguards Beijing Olympics. URL: [http://english.cas.cn/Ne/CASE/200808/t20080815\\_18764.shtml](http://english.cas.cn/Ne/CASE/200808/t20080815_18764.shtml).
- [CR12] Filipe Coelho e Cristina Ribeiro. Image abstraction in crossmedia retrieval for text illustration. In *Proceedings of the 34th European conference on Advances in Information Retrieval*, ECIR’12, pages 329–339, Berlin, Heidelberg, 2012. Springer-Verlag. URL: [http://dx.doi.org/10.1007/978-3-642-28997-2\\_28](http://dx.doi.org/10.1007/978-3-642-28997-2_28), doi:10.1007/978-3-642-28997-2\_28.

## REFERÊNCIAS

- [CSP10] Rama Chellappa, Pawan Sinha e P. Jonathon Phillips. Face Recognition by Computers and Humans. *Computer*, 43(2):46–55, February 2010. URL: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=5410708>, doi:10.1109/MC.2010.37.
- [EC97] Kamran Etemad e Rama Chellappa. Discriminant analysis for recognition of human face images. *Journal of the Optical Society of America A*, 14(8):1724, August 1997. URL: <http://josaa.osa.org/abstract.cfm?URI=josaa-14-8-1724>, doi:10.1364/JOSAA.14.001724.
- [GQP10] Patrick J Grother, George W Quinn e P Jonathon Phillips. Report on the Evaluation of 2D Still-Image Face Recognition Algorithms. *Nat'l Inst. of Standards and Technology interagency/internal report (NISTIR)*, 7709, 2010.
- [Int] Intel Corporation. What Happens In An Internet Minute? URL: <http://www.intel.com/content/www/us/en/communications/internet-minute-infographic.html>.
- [its] itseez - OpenCV Developers Team. ABOUT | OpenCV. URL: <http://opencv.org/about.html>.
- [Kan73] Takeo Kanade. *Computer recognition of human faces*. Birkhauser, Basel, 1973.
- [KCWI12] Jan Eric Kyprianidis, John Collomosse, Tinghuai Wang e Tobias Isenberg. State of the ‘Art’: A Taxonomy of Artistic Stylization Techniques for Images and Video. *IEEE Transactions on Visualization and Computer Graphics*, 2012. URL: <http://www.kyprianidis.com/p/tvcg2012/>, doi:10.1109/TVCG.2012.160.
- [KKD09] Jan Eric Kyprianidis, Henry Kang e Jürgen Döllner. Image and Video Abstraction by Anisotropic Kuwahara Filtering. *Computer Graphics Forum*, 28(7):1955–1963, October 2009. URL: <http://doi.wiley.com/10.1111/j.1467-8659.2009.01574.x>, doi:10.1111/j.1467-8659.2009.01574.x.
- [KL08] Henry Kang e Seungyong Lee. Shape-simplifying Image Abstraction. *Computer Graphics Forum*, 27(7):1773–1780, October 2008. URL: <http://doi.wiley.com/10.1111/j.1467-8659.2008.01322.x>, doi:10.1111/j.1467-8659.2008.01322.x.
- [KS90] M. Kirby e L. Sirovich. Application of the Karhunen-Loeve procedure for the characterization of human faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1):103–108, 1990. URL: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=41390](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=41390), doi:10.1109/34.41390.
- [LJ11] SZ Li e AK Jain. *Handbook of face recognition*. Springer, 2011. URL: <http://onlinelibrary.wiley.com/doi/10.1002/cbdv.200490137/abstract><http://books.google.com/books?hl=en&lr=&id=KSXwPmoqGWYC&oi=fnd&pg=PR3&q=Handbook+of+Face+Recognition&ots=lq4dqu5FWf&sig=pEu82EQkd8eougSZDpY0r5pCYLQ>.
- [LSDJ06] MS Lew, N Sebe, C Djeraba e R Jain. Content-based multimedia information retrieval: State of the art and challenges. *ACM Transactions on Multimedia*, 2(1):1–19, 2006. URL: <http://dl.acm.org/citation.cfm?id=1126005>.

## REFERÊNCIAS

- [Mina] Ministério da Administração Interna. Passaporte Electrónico Português. URL: <http://www.pep.pt/estatisticas.html>.
- [Minb] Ministério da Administração Interna. RAPID - Reconhecimento Automático de Passageiros Identificados Documentalmente. URL: <http://www.rapid.sef.pt/>.
- [PPC07] Giuseppe Papari, Nicolai Petkov e Patrizio Campisi. Artistic Edge and Corner Enhancing Smoothing. *IEEE Transactions on Image Processing*, 16(10):2449–2462, October 2007. doi:10.1109/TIP.2007.903912.
- [PRR00] P.J. Phillips, S.A. Rizvi e P.J. Rauss. The FERET evaluation methodology for face-recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1090–1104, 2000. URL: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=879790>, doi:10.1109/34.879790.
- [PSO07] PJ Phillips, WT Scruggs e AJ O’Toole. FRVT 2006 and ICE 2006 large-scale results, 2007. *National Institute of Standards and Technology*, (March), 2007. URL: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:FRVT+2006+and+ICE+2006+Large-Scale+Results#1>.
- [PWHR98] P.Jonathon Phillips, Harry Wechsler, Jeffery Huang e Patrick J. Rauss. The FERET database and evaluation procedure for face-recognition algorithms. *Image and Vision Computing*, 16(5):295–306, April 1998. URL: [http://dx.doi.org/10.1016/S0262-8856\(97\)00070-X](http://dx.doi.org/10.1016/S0262-8856(97)00070-X), doi:10.1016/S0262-8856(97)00070-X.
- [TM98] C. Tomasi e R. Manduchi. Bilateral filtering for gray and color images. In *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, pages 839–846. Narosa Publishing House, 1998. URL: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=710815>, doi:10.1109/ICCV.1998.710815.
- [TP91] Matthew Turk e Alex Pentland. Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, January 1991. URL: <http://www.mitpressjournals.org/doi/abs/10.1162/jocn.1991.3.1.71>, doi:10.1162/jocn.1991.3.1.71.
- [VB] Vision-Box. Vision-Box. URL: <http://www.vision-box.com/>.
- [VJ04] Paul Viola e Michael J. Jones. Robust Real-Time Face Detection. *International Journal of Computer Vision*, 57(2):137–154, May 2004. URL: <http://link.springer.com/10.1023/B:VISI.0000013087.49260.fb>, doi:10.1023/B:VISI.0000013087.49260.fb.
- [WOG06] Holger Winnemöller, Sven C. Olsen e Bruce Gooch. Real-time video abstraction. *ACM Transactions on Graphics*, 25(3):1221, July 2006. URL: <http://dl.acm.org/citation.cfm?id=1141911.1142018>, doi:10.1145/1141911.1142018.
- [YKMA02] Ming-hsuan Yang, David J Kriegman, Senior Member e Narendra Ahuja. Detecting faces in images: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(1):34–58, 2002. URL: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=982883>, doi:10.1109/34.982883.

## REFERÊNCIAS

- [ZCK98] W. Zhao, R. Chellappa e A. Krishnaswamy. Discriminant analysis of principal components for face recognition. In *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*, pages 336–341. IEEE Comput. Soc, 1998. URL: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=670971](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=670971), doi:10.1109/AFGR.1998.670971.
- [ZCPR03] W. Zhao, R. Chellappa, P. J. Phillips e A. Rosenfeld. Face recognition: A Literature Survey. *ACM Computing Surveys*, 35(4):399–458, December 2003. URL: <http://dl.acm.org/citation.cfm?id=954339.954342>, doi:10.1145/954339.954342.