

Visage

Impacto dos Filtros de Abstração no Reconhecimento Facial em Imagens

Pedro Pontes¹, Filipe Coelho^{1,3} e Cristina Ribeiro^{2,3}

¹ Laboratório SAPO/U.Porto
ei08091@fe.up.pt

² INESC Tecnologia e Ciência
filipe.coelho@fe.up.pt

³ Departamento de Engenharia Informática - Universidade do Porto
mrc@fe.up.pt

Resumo O reconhecimento facial em imagens constitui uma área de investigação em aberto, principalmente se considerarmos situações de captura de imagens não controladas. Os filtros de abstração atuam como ferramentas de remoção de informação redundante existente nas imagens. Foi desenvolvido um sistema de reconhecimento facial de personalidades, baseado em código aberto, onde é utilizada a abstração de imagens juntamente com tarefas de pré-processamento paralelas, de forma a analisar o seu impacto no processo de reconhecimento. A avaliação foi efetuada com recurso à coleção de imagens *Labeled Faces in the Wild*, sob duas perspetivas, *Closed-set Identification* e *Image Retrieval*, e utilizando nove cadeias de pré-processamento de imagens distintas. Os resultados demonstram que a aplicação de filtros de abstração no processo de reconhecimento resulta no compromisso entre a diminuição dos requisitos de armazenamento das imagens e a ligeira redução da eficácia da identificação. A deteção e segmentação das faces presentes nas imagens revelou ser a etapa de pré-processamento com maior importância para um reconhecimento eficaz. O desempenho foi avaliado através dos algoritmos *Eigenfaces*, *Fisherfaces* e *Local Binary Patterns Histograms*, tendo o último revelado o melhor desempenho em termos globais.

1 Introdução

A sociedade atual caracteriza-se pelo constante fluxo de informação a que os seus indivíduos estão expostos, assistindo-se a um crescimento acelerado na produção de conteúdos multimédia, tendo em conta os meios de transmissão de informação digital. Neste contexto, torna-se importante o desenvolvimento de novas metodologias que possibilitem a categorização, organização e pesquisa por parte dos utilizadores destes conteúdos [1], representando o reconhecimento facial automático em imagens um papel importante a esse nível.

Ao longo dos últimos 20 anos o reconhecimento facial em imagens sofreu uma evolução notável. Em 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, onde se verifica se existe uma correspondência entre a pessoa representada na imagem e a identidade fornecida, se encontra praticamente resolvido, uma vez que as taxas de reconhecimento atingidas são satisfatórias para a grande maioria das aplicações [2]. Contudo, o problema de reconhecimento facial automático ainda se encontra longe de ser um problema totalmente resolvido. Em cenários onde as condições de captura de imagens não são controladas, a identificação das entidades representadas nas imagens é ainda uma área de investigação em aberto [2,3].

Por seu lado, a abstração de imagens dota os artistas de novas ferramentas de transmissão de informação, tendo como objetivo, melhorar eficácia da comunicação visual. Os filtros de abstração permitem remover informação não essencial de uma imagem, dando apenas destaque à mensagem a transmitir. A aplicação de filtros de abstração de imagens na recuperação de informação multimédia, nomeadamente no âmbito da ilustração automática de texto, tem a potencialidade de melhorar a informação retornada, assim como reduzir significativamente as necessidades de processamento e armazenamento das imagens [4].

Neste artigo, encontra-se descrito o sistema de reconhecimento facial automático em imagens Visage e a avaliação efetuada ao impacto da abstração e outras tarefas de pré-processamento utilizadas no processo de reconhecimento.

2 Sistema de Reconhecimento Facial

2.1 Visão Geral

A criação do sistema de reconhecimento facial Visage teve como principal objetivo o desenvolvimento de um sistema automático de reconhecimento facial em imagens, baseado em código aberto, que permita analisar qual o impacto da abstração de imagens e outras tarefas de pré-processamento no processo de reconhecimento. Com o sistema desenvolvido pretende-se ainda a criação de uma base sólida que permita o desenvolvimento de futuras aplicações que tirem partido do reconhecimento facial automático no seu funcionamento, como por exemplo, aplicações na área de *image retrieval* e contribuir para existência de um maior número de soluções de código aberto deste tipo. De seguida encontram-se descritas as principais etapas de processamento no sistema.

2.2 Pré-processamento

A primeira tarefa desempenhada pelo sistema criado consiste no pré-processamento das imagens. Para esse fim foram desenvolvidas as seguintes sub-tarefas de pré-processamento:

1. **Deteção e Corte da Face:** Tem como objetivo determinar a zona da imagem onde se encontra a face a identificar de forma a efetuar a segmentação da face da restante imagem. A sua implementação teve por base o uso de um *boosted cascade classifier* treinado especialmente para a deteção de faces,

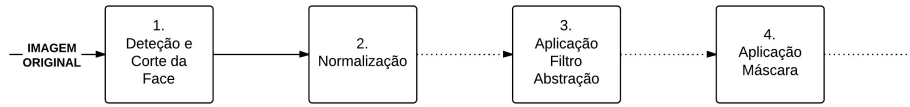


Figura 1: Tarefas de pré-processamento utilizadas na avaliação do sistema Visage.

disponível na biblioteca *Open Source Computer Vision (OpenCV)* baseado nos trabalhos de Viola e Jones [5] e posteriores melhorias introduzidas por Lienhart *et al.* [6].

2. **Normalização:** Permite efetuar a normalização do contraste da imagem recorrendo a três técnicas distintas. A primeira é designada *Contrast Stretching* e consiste em alargar a gama de valores utilizados para mapear a intensidade da imagem através de um mapeamento linear dos valores originais para uma gama de intensidade entre 0 e 255. A segunda técnica, Equalização do Histograma, consiste num mapeamento do histograma da imagem utilizando a função de distribuição acumulada (*cumulative distribution function*) de forma a que o histograma resultante se aproxime de uma distribuição mais alargada e idealmente de intensidade uniforme [7]. A última técnica, *Contrast limited adaptive histogram equalization (CLAHE)*, procura ultrapassar as limitações da técnica de equalização do histograma ao ter em consideração a intensidade de um conjunto de píxeis e dos seus vizinhos para a normalização em vez de considerar toda a imagem. Para além disso, esta última abordagem considera ainda um limite para o qual a equalização é realizada, evitando assim um mapeamento demasiado agressivo [8]. Todas as técnicas de normalização foram integradas no sistema Visage utilizando como base a biblioteca *OpenCV*.
3. **Aplicação de Filtro de Abstração:** Efetua a abstração da imagem com recurso a um dos três filtros de abstração disponíveis, gaussiano, bilateral e Kuwahara anisotrópico. A aplicação de um filtro gaussiano é a uma técnica tipicamente aplicada no processamento de imagem e computação gráfica para suavizar imagens [9] e corresponde ao menor nível de abstração dos três filtros utilizados. O filtro bilateral foi originalmente proposto por Tomasi e Manduchi [10] e caracteriza-se por suavizar as zonas com menor contraste da imagem enquanto preserva os limites de maior contraste. O filtro Kuwahara anisotrópico é uma generalização do filtro Kuwahara original que remove alguns dos artefactos originados na aplicação do filtro original através da adaptação da escala e orientação do filtro à estrutura local da imagem [11]. Os filtros gaussiano e bilateral, foram utilizados com recurso à implementação existente na biblioteca *OpenCV* e o filtro Kuwahara anisotrópico, recorrendo à implementação efetuada por Kyprianidis *et al.* [11] do mesmo filtro.
4. **Aplicação de Máscara:** Aplicação de uma máscara elíptica sobre a imagem com o objetivo de remoção do fundo ainda existente após a deteção e segmentação da face e eventuais tarefas de pré-processamento realizadas.

As sub-tarefas 2, 3 e 4 são opcionais e independentes entre si, existindo assim a possibilidade de não aplicar uma das sub-tarefas sem prejudicar as restantes.

2.3 Treino e Identificação

À fase de pré-processamento seguem-se as fases de treino e identificação. Na fase de treino é fornecido um conjunto de imagens anotadas ao sistema e pretende-se que este crie uma base de conhecimento que permita proceder à identificação de novas imagens na fase de identificação. A fase de identificação corresponde à utilização do sistema com o intuito de identificar novas imagens. Ambas as fases foram implementadas com recurso ao módulo de reconhecimento facial *FaceRecognizer*, da biblioteca *OpenCV*, que permite o treino de um modelo de reconhecimento facial com os algoritmos *Eigenfaces* [12], *Fisherfaces* [13,14] e *Local Binary Patterns Histograms (LBPH)* [15]. O módulo *FaceRecognizer* foi ainda estendido de forma a que para uma dada imagem fornecida ao sistema seja devolvida uma lista ordenada das n entidades mais parecidas com essa imagem e não apenas a entidade mais parecida.

3 Avaliação

3.1 Conjuntos de Teste

A coleção *Labeled Faces in the Wild (LFW)* [16] é uma base de dados fotográfica concebida para o estudo do problema de reconhecimento facial, particularmente em situações onde as condições de captura das imagens não possuem restrições. Nesse sentido, as imagens que constituem a galeria caracterizam-se por possuir uma grande variabilidade, nomeadamente ao nível da pose, expressão, iluminação, etnia, idade, género, vestuário e qualidade da câmara com a qual foi efetuada a captura [16]. Nesta biblioteca encontram-se representados 5749 indivíduos, ilustrados por 13233 imagens, no entanto apenas 1680 indivíduos possuem duas ou mais imagens na biblioteca. A cada imagem encontra-se também associada uma anotação textual, relativa ao nome da pessoa representada. Na avaliação de desempenho aqui reportada foi utilizada uma versão pré-alinhada da biblioteca LFW, designada LFW-a [17], de forma a remover um importante fator de variação de desempenho, mantendo a análise efetuada centrada no impacto dos filtros de abstração e restantes tarefas de pré-processamento utilizadas.

As avaliações de desempenho efetuadas no âmbito deste projeto, não se enquadram no paradigma de *pair matching* para o qual a biblioteca LFW foi originalmente criada, pelo que se tornou necessário o desenvolvimento de novos conjuntos de teste, neste caso foram criados 4 conjuntos. Cada um dos conjuntos possui um total de 1180 imagens, correspondentes a amostras biométricas de 59 indivíduos (20 imagens por indivíduo). As 20 imagens de cada indivíduo incluídas em cada um dos conjuntos de testes variam conforme o conjunto e foram selecionadas aleatoriamente. Algumas imagens encontram-se presentes nos quatro conjuntos de teste criados, enquanto outras apenas estão presentes apenas num conjunto, fazendo um total de 1592 imagens diferentes.

Tabela 1: Galerias criadas após pré-processamento.

Designação	1.Deteção e Corte	2.Normalização	3.Filtro Abstração	4.Máscara
Original	-	-	-	-
Cropped	Sim	-	-	-
Masked	Sim	-	-	Sim
Normalized	Sim	<i>Contrast Stretching</i>	-	Sim
Equalized	Sim	Equa. Histograma	-	Sim
Clahe	Sim	<i>CLAHE</i>	-	Sim
Bilateral	Sim	Equa. Histograma	Bilateral	Sim
Gaussian	Sim	Equa. Histograma	Gaussian	Sim
AKF	Sim	Equa. Histograma	Kuwahara Anisotropico	Sim

Para cada conjunto foram ainda criados dois sub-conjuntos de dados. O primeiro conjunto, designado galeria de treino (\mathcal{G}), é constituído por 80% das imagens de cada pessoa, as quais são utilizadas para o treino do sistema de reconhecimento facial. O segundo conjunto, designado provas (\mathcal{P}), contém os restantes 20% das imagens de cada pessoa, as quais são utilizadas para a avaliação do desempenho do reconhecimento. As imagens presentes nos sub-conjuntos \mathcal{G} e \mathcal{P} diferem de forma aleatória entre os quatro conjuntos criados, sendo que uma imagem pode pertencer à galeria de treino de um conjunto e às provas de outro.

3.2 Galerias pré-processadas

Na avaliação efetuada foram utilizadas quatro tarefas de pré-processamento sobre as imagens a reconhecer, sendo que para uma tarefa de pré-processamento podem existir várias opções disponíveis, como por exemplo para a abstração das imagens, onde três filtros diferentes podem ser utilizados. Foram criadas nove galerias de imagens, para a análise do impacto das diversas etapas de pré-processamento no reconhecimento facial. As galerias possuem todas as imagens contidas nos conjuntos de teste definidos na Secção 3.1, diferenciando-se pelos passos de pré-processamento a que as imagens foram sujeitas. Na Tabela 1 encontram-se sintetizadas as nove galerias criadas e as etapas de pré-processamento aplicadas em cada galeria. A Figura 2 apresenta um exemplo do estado da mesma imagem nas nove galerias pré-processadas.

3.3 Avaliação *Closed-set Identification*

O paradigma *closed-set identification* é um método padrão de avaliação do desempenho de sistemas de reconhecimento facial automático, tendo sido utilizado, por exemplo, nas avaliações FERET [18] e FRVT [19]. Este paradigma corresponde a um sub-problema de identificação em que uma prova é apresentada ao sistema pretendendo-se que este devolva a identidade da pessoa presente na imagem. Neste caso particular do problema de identificação, todas as provas apresentadas possuem uma correspondência na galeria, em oposição ao caso geral de identificação, no qual pode ou não haver uma correspondência.

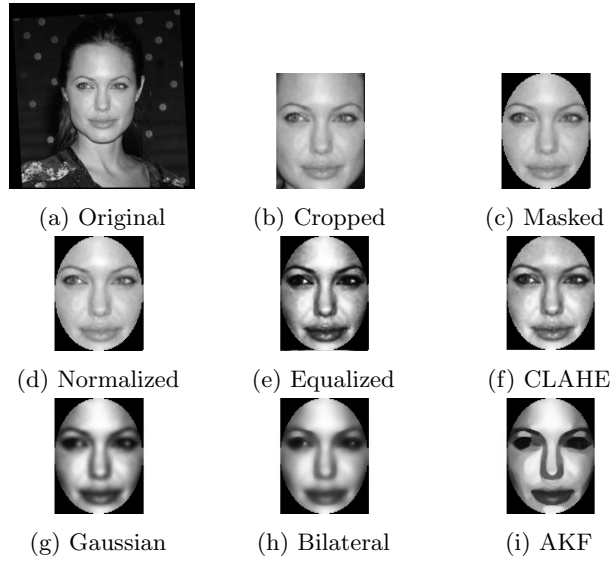


Figura 2: Variação da mesma imagem nas várias galerias pré-processadas.

Para cada conjunto de teste descrito na Secção 3.1, seja $\mathcal{G} = \{g_1, \dots, g_N\}$ a sua galeria de treino, e seja $\mathcal{P} = \{p_1, \dots, p_N\}$ o conjunto das suas provas. Quando uma prova p_j é apresentada ao sistema, essa prova é comparada com cada amostra biométrica g_i da galeria, resultando dessa comparação a respetiva medida de similaridade (*similarity score*), s_{ij} . Esta medida é designada de *match score*, caso g_i e p_j sejam amostras da mesma pessoa, caso não o sejam é designado de *nonmatch score*. Quanto menor a medida de similaridade, maior é a probabilidade das imagens comparadas pertencerem à mesma pessoa, sendo que o melhor valor possível para a medida de similaridade de um *match score* é zero. Na identificação de $p_j \in \mathcal{P}$, em primeiro lugar são calculados os índices de similaridade para todas as amostras na galeria \mathcal{G} , sendo posteriormente ordenados os seus resultados. O *rank* de p_j , r_{p_j} , é igual a n , se o seu *match score* corresponde à n -ésima menor medida de similaridade.

Na avaliação de um conjunto de teste é calculado o *rank* de cada uma das suas provas. Considerando que $|C(n)|$ corresponde ao número de provas com *rank* n , ou menor do que n , e que $|\mathcal{P}|$ corresponde ao número de provas existentes em \mathcal{P} , a taxa de identificação para o *rank* n , $T_I(n)$, corresponde à fração de provas com *rank* n ou menor do que n , ou seja:

$$T_I(n) = \frac{|C(n)|}{|\mathcal{P}|} \times 100 \quad (1)$$

A $T_I(n)$ é calculada para cada um dos *ranks* de 1 a 30. Cada galeria é avaliada com os três algoritmos disponíveis, nos quatro conjuntos de teste sendo posteriormente calculada a média dos resultados obtidos para essa galeria. Os

resultados apresentados de seguida correspondem à média dos resultados obtidos e encontram-se aqui reportados num conjunto de três experiências, cada uma dedicada a analisar a variação do comportamento do sistema dado um problema específico.

Experiência 1: Impacto Segmentação das Faces. A primeira experiência realizada visa analisar o impacto da deteção e posterior segmentação das faces presentes numa imagem. Nesta experiência foram utilizadas as galerias de imagens Original, Cropped e Masked. A primeira é constituída pelas imagens originais da biblioteca LFW-a e permite estabelecer uma base de comparação entre as imagens segmentadas e as imagens originais. A galeria Cropped é composta por um conjunto de imagens onde as faces foram detetadas e segmentadas pelo detetor facial implementado no sistema Visage. A terceira e última galeria criada possui as mesmas imagens da galeria Cropped, sobre as quais foi posteriormente aplicada uma máscara elíptica de modo a aumentar a área de fundo removida das imagens recortadas. Os resultados obtidos no âmbito desta experiência encontram-se sintetizados na Tabela 2.

Tabela 2: Média dos resultados obtidos nos 4 conjuntos de teste para as galerias Original, Cropped e Masked na Experiência 1 (melhores a negrito).

Rank	<i>Eigenfaces</i>			<i>Fisherfaces</i>			<i>LBPH</i>		
	Orig.	Crop.	Mask.	Orig.	Crop.	Mask	Orig.	Crop.	Mask
1	12.9%	25.9%	28.1%	30.1%	44.2%	43.6%	7.6%	47.3%	46.8%
5	31.9%	50.0%	52.8%	58.8%	66.3%	66.5%	21.6%	70.8%	71.9%
10	47.6%	64.0%	65.8%	71.3%	76.6%	75.3%	34.1%	79.8%	79.9%
15	57.7%	72.9%	75.0%	80.6%	81.7%	81.0%	43.9%	87.1%	85.8%
20	65.2%	81.3%	81.5%	86.4%	86.4%	84.4%	53.8%	90.4%	90.2%
25	72.0%	86.2%	86.1%	90.2%	90.2%	88.6%	63.5%	92.4%	93.0%
30	78.4%	89.5%	89.3%	92.2%	92.5%	91.0%	92.3%	94.0%	95.0%

Experiência 2: Impacto da Normalização do Contraste. A aquisição de imagens em condições não controladas pode resultar na obtenção de fotografias com um contraste reduzido e níveis de iluminação muito distintos. De forma a ultrapassar este problema, a normalização do contraste de uma imagem é uma tarefa comum nos sistemas de reconhecimento facial. Nesta experiência pretendemos analisar qual o impacto da normalização do contraste de uma imagem nos diferentes algoritmos avaliados, assim como determinar qual a melhor estratégia de normalização, de entre três técnicas implementadas, através da análise do desempenho do sistema nas galerias Normalized, onde foi aplicada a técnica *Contrast Stretching*; Equalized, onde foi efetuada a Equalização do Histograma; e Clahe, onde foi utilizada a técnica homónima. Os resultados obtidos no âmbito desta experiência encontram-se sintetizados na Tabela 3.

Tabela 3: Média dos resultados obtidos nos 4 conjuntos de teste para as galerias Normalized, Equalized e Clahe na Experiência 2 (melhores a negrito).

Rank	<i>Eigenfaces</i>			<i>Fisherfaces</i>			<i>LBPH</i>		
	Norm.	Equa.	Clahe.	Norm.	Equa.	Clahe	Norm.	Equa.	Clahe
1	30.0%	43.0%	40.9%	42.2%	47.3%	45.8%	45.7%	49.1%	46.1%
5	56.7%	68.3%	68.0%	67.5%	67.5%	66.8%	70.6%	71.8%	68.3%
10	70.4%	77.9%	80.3%	75.1%	77.5%	77.7%	80.9%	80.3%	79.5%
15	79.7%	84.3%	86.8%	81.1%	83.7%	83.3%	87.3%	85.6%	86.4%
20	85.0%	89.3%	90.6%	85.3%	87.1%	87.7%	91.0%	89.3%	90.8%
25	89.4%	91.4%	93.4%	88.5%	89.1%	89.7%	94.0%	92.4%	93.3%
30	92.3%	94.0%	95.0%	91.3%	91.1%	91.5%	95.7%	94.2%	95.6%

Experiência 3: Impacto da Abstração de Imagens. A terceira e última experiência realizada visa analisar qual o impacto da utilização de filtros de abstração no processo de reconhecimento facial em imagens. A análise foi efetuada com recurso a três filtros de abstração, gaussiano, bilateral e Kuwahara anisotrópico, os quais correspondem a diferentes técnicas e níveis de abstração, sendo o filtro gaussiano o filtro com um menor impacto na imagem original e o filtro Kuwahara anisotrópico aquele que produz um maior grau de abstração. Os resultados obtidos no âmbito desta experiência encontram-se sintetizados na Tabela 3.

Tabela 4: Média dos resultados obtidos nos 4 conjuntos de teste para as galerias Gaussian, Bilateral e AKF na Experiência 3 (melhores a negrito).

Rank	<i>Eigenfaces</i>			<i>Fisherfaces</i>			<i>LBPH</i>		
	Gaus.	Bila.	AKF.	Gaus.	Bila.	AKF	Gaus.	Bila.	AKF
1	43.4%	42.7%	41.0%	44.4%	44.2%	42.7%	47.5%	42.3%	31.5%
5	68.5%	67.9%	66.1%	64.8%	67.4%	66.4%	69.4%	66.3%	56.6%
10	78.3%	77.7%	76.4%	75.6%	76.8%	77.9%	80.0%	78.4%	67.3%
15	85.0%	85.1%	83.4%	81.7%	81.1%	83.7%	86.3%	83.3%	74.1%
20	89.4%	89.7%	88.5%	85.0%	86.0%	86.9%	90.7%	86.6%	81.4%
25	91.7%	91.8%	91.2%	88.7%	89.0%	91.4%	93.3%	89.0%	85.0%
30	94.4%	94.5%	94.3%	91.5%	91.3%	94.2%	95.7%	92.1%	87.9%

3.4 Avaliação *Image Retrieval*

Em recuperação de informação designa-se *precisão* a fração de resultados relevantes do total de resultados retornados por uma pesquisa. No sistema de reconhecimento facial criado, para uma imagem fornecida é apresentada uma lista de possíveis entidades que se encontram representadas nessa imagem, em que o primeiro resultado representa a entidade com maior probabilidade de estar presente na imagem e em que cada entidade aparece uma única vez na lista de resultados. Caso se pretenda efetuar a identificação de uma personalidade existe

Tabela 5: Resultados da Média da Precisão na Galeria (melhores a negrito).

Galeria	$MPG_{Eigenfaces}$	$MPG_{Fisherfaces}$	MPG_{LBPH}
Original	23.1%	43.4%	16.0%
Cropped	37.7%	54.8%	58.2%
Masked	40.0%	54.0%	58.3%
Normalized	42.4%	53.5%	57.2%
Equalized	55.2%	57.3%	59.6%
CLAHE	53.6%	55.9%	56.5%
Gaussian	55.3%	54.2%	58.1%
Bilateral	54.6%	54.7%	53.8%
AKF	52.8%	53.9%	43.0%

apenas um resultado relevante na lista de resultados obtida. Ao analisar o desempenho do sistema torna-se importante verificar a posição em que surgem os resultados relevantes e não apenas a fração de resultados relevantes analisada pela *precisão*.

Para avaliação do desempenho do sistema do ponto de vista de *image retrieval* foi criada a medida *precisão na galeria* (PG). A *precisão na galeria* resulta da adaptação da precisão ao caso particular do sistema de reconhecimento facial automático desenvolvido. Considerando a definição de *rank* de uma prova apresentada na Secção 3.3, seja \mathcal{P} o conjunto de provas da galeria, $p_j \in \mathcal{P}$ e r_{p_j} o *rank* da prova p_j . A *precisão na galeria* de p_j , PG_{p_j} , é definida como:

$$PG_{p_j} = \frac{1}{r_{p_j}} \times 100 \quad (2)$$

A *precisão na galeria* é uma precisão no sentido em que é uma razão entre número de documentos relevantes (aqui sempre 1) e número de documentos na resposta. Nesta avaliação o que consideramos é uma resposta de tamanho variável e igual ao número de amostras que é necessário considerar até incluir a relevante. A *precisão na galeria* é maior quanto mais cedo o nome da pessoa a identificar aparece na lista de resultados possíveis. Por outro lado, a distância entre precisões na galeria sucessivas é maior quanto maior forem os valores de precisão. Como tal, é possível efetuar uma análise do desempenho do sistema com ênfase nos resultados de topo, ao mesmo tempo que se diferenciam os valores de precisão mais elevados, os quais representam os resultados mais significativos para os utilizadores do sistema.

A análise do desempenho de um algoritmo é efetuada com recurso à análise do seu desempenho para um conjunto de provas. Seja n o número de provas existentes em \mathcal{P} , a *média da precisão na galeria* (MPG) de um algoritmo é calculada através da média da *precisão na galeria* obtida para cada prova e pode ser definida como:

$$MPG_{algoritmo} = \frac{\sum_{i=0}^n PG_{p_i}}{n} \quad (3)$$

A *média da precisão na galeria* para as nove galerias avaliadas encontra-se reportada na Tabela 5.

4 Requisitos Armazenamento

Um sistema de reconhecimento facial automático pode ser analisado do ponto de vista do seu desempenho no reconhecimento, no entanto, as necessidades de processamento e armazenamento das imagens utilizadas por este tipo de sistemas devem também ser objeto de análise.

Na Tabela 6 encontram-se sintetizadas as necessidades de armazenamento das imagens que compõem as galerias utilizadas na avaliação do sistema Visage. Todas as galerias pré-processadas reduzem em pelo menos 81% as necessidades de armazenamento necessárias. A este nível as imagens da galeria AKF, abstraídas com recurso ao filtro Kuwahara anisotrópico, destacam-se uma vez que são em média 89% inferiores às imagens originais. A redução do espaço de armazenamento com recurso à abstração permite a criação de galerias mais fáceis e rápidas de processar pelo sistema, ao mesmo tempo que reduz da quantidade de dados transferidos na utilização do sistema Visage em cooperação com eventuais clientes web ou aplicação móveis.

Tabela 6: Necessidades de armazenamento galerias pré-processadas.

Galeria	Total(MB)	Médio (KB)	Máximo (KB)	Mínimo (KB)	Diferença
Original	125.5	81	114	47	-
Cropped	20.9	13	18	9	-83.4%
Masked	19.4	12	16	9	-84.5%
Normalized	20.7	13	16	10	-83.5%
Equalized	23.7	15	18	12	-81.1%
CLAHE	22.9	15	18	12	-81.7%
Gaussian	22.4	14	17	12	-82.2%
Bilateral	21.7	14	16	11	-82.7%
AKF	13.7	9	11	6	-89.1%

5 Discussão e Conclusão

Neste artigo avaliamos o sistema de reconhecimento facial Visage ao nível do seu desempenho no reconhecimento, e também ao nível dos requisitos de armazenamento necessários utilizando nove cadeias de pré-processamento de imagens.

As avaliações efetuadas permitem concluir que a etapa de pré-processamento que permite um maior aumento no número de indivíduos corretamente reconhecidos é a deteção e segmentação das faces presentes numa imagem. Através da comparação do desempenho das galerias Original, Cropped e Masked é ainda possível concluir que a segmentação da face da restante imagem produz um maior impacto no algoritmo *LBPH*, seguido pelo algoritmo *Fisherfaces* e em último lugar pelo algoritmo *Eigenfaces*. Finalmente, a este nível é ainda possível

destacar que não se registou uma diferença significativa no desempenho do sistema entre as imagens com máscara e as imagens apenas recortadas. A aplicação de uma máscara constitui, no entanto, uma técnica tipicamente aplicada para a remoção de uma maior quantidade de fundo das imagens [18,15,20], uma vez que permite obter uma maior confiança nos resultados obtidos ao garantir que o reconhecimento efetuado não tira partido dos elementos existentes no fundo da imagens, mas sim das características faciais representadas [20].

Ao nível da normalização do contraste das imagens verificou-se que a aplicação desta etapa de pré-processamento permite uma melhoria no número de faces corretamente reconhecidas, com particular ênfase no algoritmo *Eigenfaces*, onde se registou uma melhoria de 15,2% na comparação da *média da precisão na galeria* das galerias Equalized e Masked. Na comparação das três técnicas de normalização do contraste, a equalização simples do histograma e a técnica *CLAHE* foram as que revelaram melhores resultados para os algoritmos *Eigenfaces* e *Fisherfaces*. No caso algoritmo *LBPH*, a técnica de *Contrast Stretching* e a equalização simples do histograma revelaram os melhores resultados. Contudo, a equalização do histograma demonstrou um desempenho ligeiramente superior na taxa de identificação para os primeiros *ranks*, tendo sido por isso utilizada como técnica de normalização na cadeia de pré-processamento das imagens abstraídas. A reduzida diferença no desempenho das galerias Normalized, Equalized e CLAHE justifica um estudo mais aprofundado das técnicas de normalização do contraste em trabalho futuro.

Por último, a avaliação efetuada à utilização de abstração de imagens no âmbito do processo de reconhecimento facial automático revela a existência de um compromisso entre a redução das necessidades de armazenamento das imagens abstraídas e uma ligeira diminuição do desempenho do reconhecimento efetuado. A este nível destaca-se a utilização do filtro Kuwahara anisotrópico para a abstração das imagens, o qual permite uma redução de 89% do espaço de armazenamento, implicando no entanto uma diminuição de cerca de 6% na *média da precisão da galeria* quando comparadas as galerias AKF e Equalized. A aplicação da abstração com recurso aos filtros gaussiano e bilateral demonstrou um menor impacto no desempenho do reconhecimento, mas também uma menor redução nos requisitos de armazenamento das respetivas galerias. Este fator pode ser explicado devido à maior semelhança entre as imagens abstraídas com recurso a estes filtros e as imagens não abstraídas, devido ao menor grau de abstração por eles aplicado. A combinação da galeria Gaussian com a utilização do algoritmo *LBPH* obteve a taxa de identificação mais elevada das galerias abstraídas, sendo esta, no entanto, ligeiramente inferior às taxas de identificação mais elevadas das galerias utilizadas na Experiência 2.

Referências

1. Datta, R., Joshi, D., Li, J., Wang, J.Z.: Image retrieval: Ideas, Influences and Trends of the New Age. *ACM Computing Surveys* **40**(2) (April 2008) 1–60
2. Li, S., Jain, A.: *Handbook of face recognition*. Springer (2011)

3. Pinto, N., DiCarlo, J., Cox, D.: How far can you get with a modern face recognition test set using only simple features? In: IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL (2009) 2591 – 2598
4. Coelho, F., Ribeiro, C.: Image abstraction in crossmedia retrieval for text illustration. In: Proceedings of the 34th European conference on Advances in Information Retrieval. ECIR'12, Berlin, Heidelberg, Springer-Verlag (2012) 329–339
5. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001 **1** (2001) I-511–I-518
6. Lienhart, R., Kuranov, A., Pisarevsky, V.: Empirical analysis of detection cascades of boosted classifiers for rapid object detection. Pattern Recognition (2003)
7. Bradski, G., Kaehler, A.: Learning OpenCV: Computer vision with the OpenCV library. O'Reilly Media, Inc. (2008)
8. Reza, A.M.: Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement. The Journal of VLSI Signal Processing-Systems for Signal, Image, and Video Technology **38**(1) (August 2004) 35–44
9. Deng, G., Cahill, L.W.: An adaptive Gaussian filter for noise reduction and edge detection. In: Nuclear Science Symposium and Medical Imaging Conference, 1993., 1993 IEEE Conference Record. (1993) 1615–1619 vol.3
10. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271), Narosa Publishing House (1998) 839–846
11. Kyprianidis, J.E., Kang, H., Döllner, J.: Image and Video Abstraction by Anisotropic Kuwahara Filtering. Computer Graphics Forum **28**(7) (October 2009) 1955–1963
12. Turk, M., Pentland, A.: Eigenfaces for Recognition. Journal of Cognitive Neuroscience **3**(1) (January 1991) 71–86
13. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence **19**(7) (July 1997) 711–720
14. Zhao, W., Chellappa, R., Krishnaswamy, A.: Discriminant analysis of principal components for face recognition. In: Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition, IEEE Comput. Soc (1998) 336–341
15. Ahonen, T., Hadid, A., Pietikäinen, M.: Face recognition with local binary patterns. Computer Vision-ECCV 2004 (2004) 469–481
16. Huang, G.B., Ramesh, M., Berg, T., Learned-miller, E.: Labeled Faces in the Wild : A Database for Studying Face Recognition in Unconstrained Environments. Technical report, University of Massachusetts, Amherst (2007)
17. Taigman, Y., Wolf, L., Hassner, T.: Multiple one-shots for utilizing class label information. In: The British Machine Vision Conference (BMVC). (Sept. 2009)
18. Phillips, P., Rizvi, S., Rauss, P.: The FERET evaluation methodology for face-recognition algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence **22**(10) (2000) 1090–1104
19. Blackburn, D.M., Bone, M., Phillips, P.J.: Face Recognition Vendor Test 2000: Evaluation Report. (February 2001)
20. Kumar, N., Berg, A.C., Belhumeur, P.N., Nayar, S.K.: Attribute and Simile Classifiers for Face Verification. In: IEEE International Conference on Computer Vision (ICCV). (Oct 2009)