

Avaliação de Desempenho de Reconhecimento Automático de Faces Utilizando as Técnicas de Análise de Componentes Principais e de *Fisherfaces*

**Rosana de Paula Lucas, Flávio Henrique Teles Vieira, Webert Brito dos Passos,
Douglas Leo Netzlaff**

Escola de Engenharia Elétrica, Mecânica e de Computação – Universidade Federal de
Goiás (UFG)

Av. Universitária, n. 1488 - Quadra 86 - Bloco A - 3º piso 74605-010 - Setor Leste
Universitário - Goiânia – GO – Brazil

rosana.rpl@hotmail.com, flavio@eeec.ufg.br, webertbrito@gmail.com,
douglas_ntz@hotmail.com

***Abstract.** In this paper, we evaluate the face recognition performance of two different techniques based on PCA (Principal Components Analysis) and Fisherfaces. These algorithms are applied to two image banks well-known in the literature, presenting different characteristics. From the comparison of the simulation results, we intend to point out which algorithm provides best face recognition performance for the two image banks under different conditions, for example, by adding noise to the images.*

***Resumo.** Neste trabalho, avaliamos o desempenho de reconhecimento de faces em imagens de dois algoritmos diferentes baseados nas técnicas de PCA (Principal Components Analysis) e de Fisherfaces. Estes algoritmos são aplicados a dois bancos de imagens conhecidos na literatura que apresentam características diferentes. A partir da comparação dos resultados de simulação, buscamos apontar qual o melhor algoritmo em termos de desempenho de reconhecimento de faces para as imagens utilizadas sob diferentes condições, como por exemplo, adicionando ruído às imagens.*

1. Introdução

A face é um elemento biométrico individualizado e de fácil obtenção, o que torna a identificação facial uma área de constantes pesquisas.

Os computadores sofrem uma série de limitações e dificuldades para realizar esta tarefa. Para os seres humanos, a identificação e o reconhecimento facial são simples, embora os seus mecanismos ainda não tenham sido totalmente explicados, sendo fonte de grande interesse para os pesquisadores da área de reconhecimento de padrões. Diversos estudos vêm sendo realizados com potencial não só para explicar os mecanismos cerebrais de reconhecimento de faces como, também, para auxiliar no reconhecimento por máquinas/computadores, muito embora tais pesquisas tenham pouca sinergia [Chellappa 1995].

Houve grande crescimento de pesquisas a partir da década de 90, buscando solucionar problemas como encontrar e isolar uma face em uma imagem, e extrair características específicas, como olhos e boca. Várias técnicas foram desenvolvidas e aplicadas, como a Análise de Componentes Principais (PCA – *Principal Components Analysis*), baseada na Transformada de Karhunen-Lòeve (KLT), o *Singular Value Decomposition* (SVD), classificadores baseados em redes neurais, a Análise de Discriminantes Lineares (LDA – *Linear Discriminant Analysis*), dentre outros [Shakhnarovich 2004].

Neste artigo, avaliamos o desempenho de reconhecimento de faces em imagens dos algoritmos PCA e de *Fisherfaces* sob diferentes fatores como número de testes, presença de ruído e variação do número de imagens disponíveis por pessoa, etc.

2. Análise de Componentes Principais

A Análise de Componentes Principais (PCA) foi criada em 1901, por Karl Pearson, matemático britânico, que buscava representar de forma mais simples dados de dimensão elevada. Em 1933, Harold Hotelling usou a técnica de Pearson em matrizes, criando o termo “Componentes Principais”. Em 1987, L. Sirovitch e M. Kirby, no trabalho “*Low-Dimensional Procedure for the Characterization of Human Faces*” [Sirovitch 1987], propuseram o uso do PCA com o objetivo de reduzir as dimensões de imagens, perdendo o mínimo possível de informações. Devido à grande redundância presente nas imagens de face, a técnica teve sucesso nesta área. Após a publicação do trabalho de Sirovitch e Kirby, uma série de outros trabalhos foram desenvolvidos, com destaque para Mathew Turk e Alex Pentland, em 1991 [Turk 1991].

Entretanto, há limitações para o uso do PCA: é necessário haver linearidade no conjunto de dados a serem manipulados, a média e a variância devem ser estatísticas suficientes e, como as maiores variâncias são destacadas, presume-se que elas representam relações significativas, enquanto variâncias menores são menos relevantes.

Apresentamos a seguir os passos do Algoritmo PCA implementado neste trabalho:

Passo 1 Subtração pela média e cálculo da matriz de covariâncias

Primeiramente, deve-se subtrair todos os elementos da imagem pela média da sua dimensão correspondente. Deste modo, a média deste novo conjunto de dados será igual a zero. Calcula-se, então, a matriz de covariâncias. Sendo um conjunto de dados de duas dimensões, a matriz terá a seguinte forma:

$$C = \begin{pmatrix} \text{cov}(x,x) & \text{cov}(x,y) \\ \text{cov}(y,x) & \text{cov}(y,y) \end{pmatrix} \quad (1)$$

Passo 2 Cálculo dos autovetores e autovalores [Turk 1991]

Calcula-se os autovetores e autovalores da matriz de covariância (equação (1)). Neste caso, teremos 2 autovetores. Por terem aparência semelhante a faces, os autovetores são denominados autofaces ou *eigenfaces* conforme mostra a Figura 1.

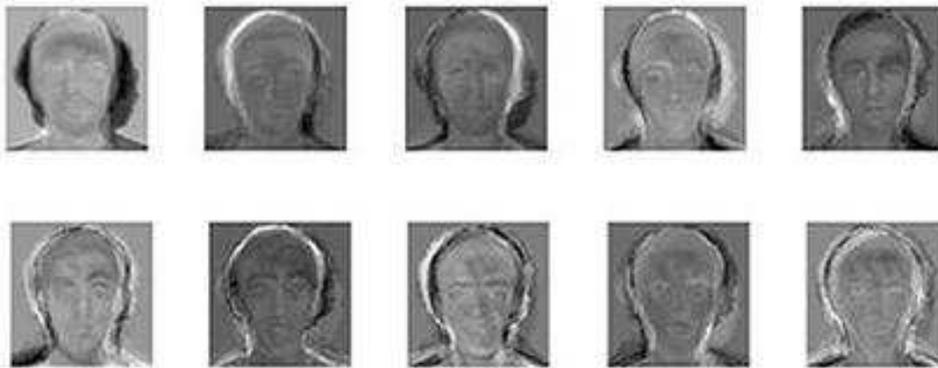


Figura 1 – Autofaces [Kruger 2004]

Passo 3 Ordenação dos componentes e redução de dimensionalidade

Uma vez calculados os autovetores, pode-se ordená-los em ordem decrescente, de acordo com os autovalores e, conseqüentemente, de importância. Com os autovetores ordenados de acordo com sua importância, podem ser ignorados os autovetores com valores mais baixos. Haverá perda de informação, mas se os autovalores forem pequenos, essa perda será aceitável e, em muitos casos, desejada, visto que muitas vezes estes componentes estarão associados a ruído.

Passo 4 Criação do Vetor de Características e Representação dos Dados

Escolhidos os autovetores que permanecerão no conjunto, deve-se organizá-los em colunas em uma matriz, tal que:

$$\text{VetorCaract} = (\text{autovet}_1 \quad \text{autovet}_2 \quad \dots \quad \text{autovet}_n) \quad (2)$$

Agora, deve-se obter o vetor de características transposto, de tal forma que cada autovetor ocupará uma linha da matriz. O mesmo deve ser feito com a matriz dos dados, subtraídos pelas médias. Assim, cada linha da matriz representará uma dimensão. Fazendo a multiplicação, temos:

$$\text{DadosFinais} = \text{VetorCaractLinha} \times \text{DadosAjustados} \quad (3)$$

sendo *Dadosajustados* os dados iniciais subtraídos da média, com cada linha representando uma dimensão.

Passo 5 Recuperação dos Dados

A recuperação total dos dados só será possível caso nenhum autovetor tenha sido descartado. Quanto maior o número de autovetores ignorados, maior será a diferença entre os dados originais e os recuperados. Reescrevendo a fórmula acima, temos que:

$$\text{DadosAjustados} = \text{VetorCaractLinha}^{-1} \times \text{DadosFinais} \quad (4)$$

Agora, basta somar as médias que foram inicialmente subtraídas para recuperar os dados:

$$\text{Dados Recuperados} = (\text{VetorCaractLinha}^{-1} \times \text{DadosFinais}) + \text{média} \quad (5)$$

Mesmo que autovetores tenham sido descartados, a fórmula acima pode ser usada. Naturalmente, caso todos os autovetores tenham sido mantidos, essa matriz apresentará os dados originais, e caso descartes tenham sido feitos, haverá perda de informação.

3. Fisherfaces

Duas das maiores dificuldades no processo de identificação facial estão relacionadas à variação de iluminação e de expressões faciais. O PCA retém estas informações, que são indesejadas e podem aumentar as taxas de erro, ao extrair as características com maior variância. Dessa forma, os componentes principais obtidos através do PCA podem não ser os mais adequados para discriminação.

Ronald A. Fisher, em 1936, propôs um critério estatístico que visava maximizar a separação entre as classes, e ao mesmo tempo minimizar o espalhamento dentro delas. Esta técnica foi denominada Análise de Discriminantes de Fisher Lineares (FLD – *Fisher's Linear Discriminant*), ou LDA (*Linear Discriminant Analysis*) e é uma técnica supervisionada, visto que necessita-se de ação externa para organizar as classes [Sharmaa 2012].

Em [Belhumeur, 1997], os autores desenvolveram um algoritmo de reconhecimento facial baseado no FLD, com boas taxas de acerto mesmo em situações com variação de iluminação e de expressões faciais. Como as classes de faces estão separadas e nomeadas, pode-se usar estas informações para maximizar a relação entre o determinante da matriz de espalhamento inter-classes com o determinante da matriz de espalhamento intra-classes:

$$\text{Critério de Fisher} = \frac{\det|S_b|}{\det|S_w|} \quad (6)$$

A matriz de espalhamento inter-classes S_b e a matriz intra-classe S_w são definidas por:

$$S_b = \sum_{i=1}^c N_i (\mu_i - \mu)(\mu_i - \mu)^T \quad (7)$$

$$S_w = \sum_{i=1}^c \sum_{x_k \in X_i} (x_k - \mu_i)(x_k - \mu_i)^T \quad (8)$$

sendo μ_i a face média de uma classe X_i e N_i o número de imagens em uma classe. Pode-se encontrar uma base W_{opt} que maximize o critério de Fisher, da seguinte forma [Belhumeur, 1997]:

$$W_{opt} = \arg \max_w \frac{|W^T S_b W|}{|W^T S_w W|} = [w_1 \quad w_2 \quad \dots \quad w_M] \quad (9)$$

em que ϖ_i é o conjunto de autovetores generalizados de S_b e S_w correspondentes aos maiores autovalores.

Neste trabalho, aplicamos o algoritmo proposto em [Belhumeur, 1997], conhecido como técnica de *Fisherfaces* em uma base diferente de imagens considerando também diferentes fatores como adição de ruído e comparando com os resultados do algoritmo PCA.

4. Simulações e Resultados

Buscando uma análise mais ampla, foram utilizados dois bancos de faces com características diferentes: *ORL face database* [Cambridge 2012] e *Yale Face Database* [Yale 2012]. O primeiro banco, *ORL face database*, é composto por imagens que foram geradas entre abril de 1992 e abril de 1994 no *Olivetti Research Laboratory* em Cambridge, Inglaterra. São 40 indivíduos, cada um com 10 imagens distintas, totalizando 400 faces. As imagens foram feitas variando a iluminação, expressões faciais (olhos abertos ou fechados, sorrindo ou não sorrindo), e detalhes faciais (presença ou ausência de óculos). Para alguns dos indivíduos as fotos foram feitas com certo intervalo de tempo. Todas as fotos foram feitas sob fundo escuro e homogêneo. Todas possuem dimensão 112x92 e são em escala de cinza.

O segundo banco de imagens, *Yale Face Database*, é formado por imagens geradas pela Universidade Yale, nos Estados Unidos. Este banco é composto por 165 faces de 15 pessoas. São 11 faces por indivíduo, uma para cada expressão facial, com as seguintes configurações: luz central, feliz, luz à esquerda, com/sem óculos, normal, luz à direita, triste, sonolento, surpreso e piscando. A Tabela 1 apresenta os números de imagens de treinamento e testes utilizadas nas nossas simulações.

Tabela 1 - Relação da quantidade de faces distribuídas por utilidade

Faces					
Origem	Indivíduos	Treinamento		Teste	
		Individual	Total	Individual	Total
Yale	15	10	150	1	15
Cambridge	40	9	360	1	40

Sob condições normais, como mostra a Figura 2, o PCA apresenta uma melhor resposta em termos de taxa de acerto de reconhecimento de face na fase de teste para as imagens de Cambridge, mesmo variando o número de simulações (testes). Contudo, com a inserção de ruído, o PCA perde sua eficácia, enquanto a técnica de *Fisherface* mantém taxas de acerto bem mais satisfatórias, como mostra a Figura 3.

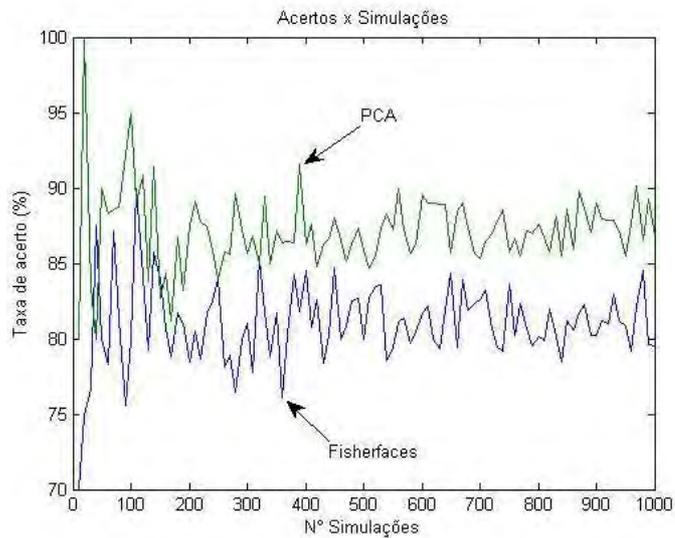


Figura 1 - PCA x Fisherfaces: taxa de acerto versus número de simulações, Cambridge

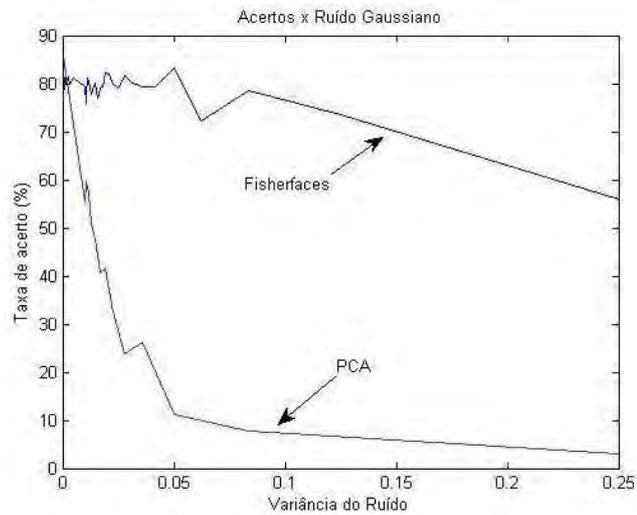


Figura 3 - PCA x Fisherfaces: Adição de ruído gaussiano, Cambridge

Para as faces de Yale percebe-se que a técnica de *Fisherfaces* apresenta desempenho superior em termos de taxa de acerto de reconhecimento de faces, como mostra a Figura 4.

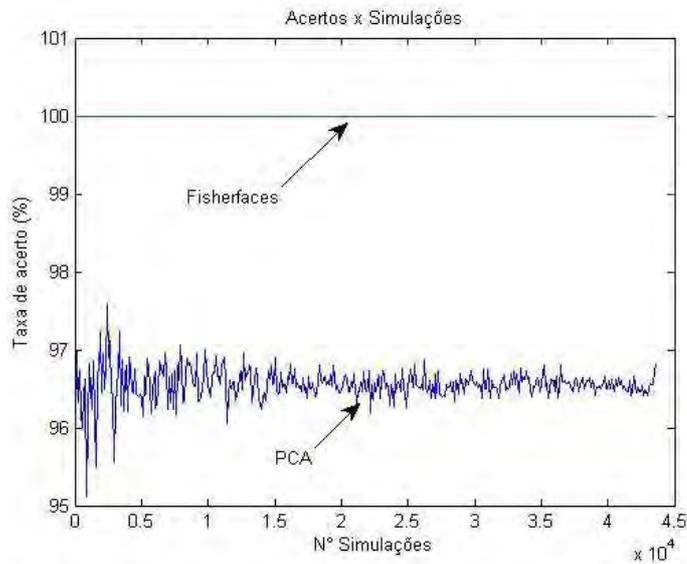


Figura 4 - PCA x Fisherfaces: taxa de acerto versus número de simulações, Yale

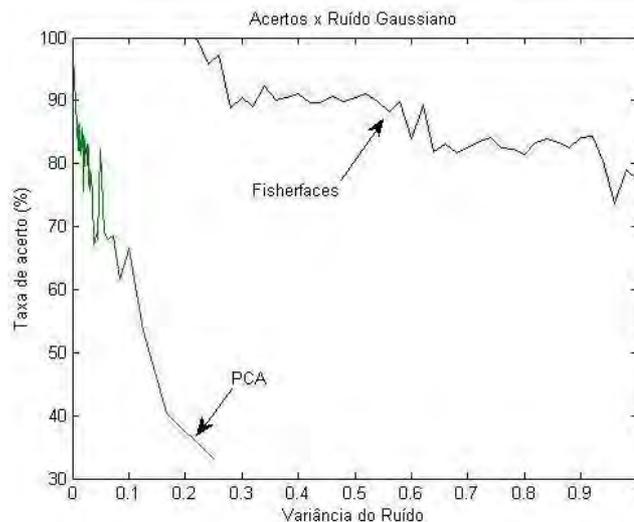


Figura 5 - PCA x Fisherfaces: Adição de ruído gaussiano, Yale

Novamente, observa-se uma maior taxa de acerto de reconhecimento de faces para o algoritmo de *Fisherfaces* quando se aumenta a variância do ruído gaussiano branco nas imagens. Notou-se também que à medida que se aumenta o número de imagens por indivíduo, o algoritmo de *Fisherfaces* obtém melhor resultado do que o algoritmo PCA.

5. Conclusões

Neste trabalho buscou-se simular e analisar dois dos algoritmos utilizados para identificação de faces, a Análise de Componentes Principais e a Análise com *Fisherfaces*.

O método de *Fisherfaces*, que é baseado no FLD, utiliza-se do PCA para possibilitar a extração de características mais discriminantes. As imagens das faces

devem ser agrupadas em classes, cada uma representando uma pessoa. Os dados são organizados de tal forma a maximizar a relação entre o espalhamento entre classes e o espalhamento intra-classes, de forma a facilitar a identificação. Os resultados obtidos nos testes mostram que, de fato, a técnica de *Fisherfaces* obtém melhores taxas de acerto em quase todas as simulações realizadas quando há presença de ruído gaussiano ou variação na iluminação.

Observou-se também variação considerável nos resultados das duas técnicas de acordo com o banco de faces utilizado. Em todos os testes realizados, o banco de Yale possibilitou maiores taxas de acerto. Comparando-se os bancos, Yale possui apenas 15 indivíduos, com 11 imagens para cada um, enquanto Cambridge possui 40, com 10 imagens para cada. Estes aspectos já explicam boa parte das diferenças nos resultados, visto que Yale possui mais informações para discriminar uma quantidade significativamente menor de indivíduos. Deve-se atentar, também, para a qualidade das imagens de cada banco. A resolução das imagens do banco de Yale é 243x320, superior a Cambridge, com resolução de 112 x 92. Yale também privilegiou fotos com variação de expressões faciais, resultando em maiores informações acerca de um indivíduo. A união de todos estes fatores pode explicar as grandes diferenças observadas nas taxas de acerto obtidas em cada banco. Evidencia-se, então, que no desenvolvimento de um sistema de identificação facial deve-se empregar atenção especial na construção do banco de faces, buscando sempre qualidade e o maior número possível de imagens e expressões faciais por pessoa.

Referências

- Cambridge University (2012): Olivetti Research Laboratory. Disponível em: http://www.cl.cam.ac.uk/Research/DTG/attarchive:pub/data/att_faces.zip. Acessado em 10/11/2012
- Chellappa, R. et al. (1995). Human and Machine Recognition of Faces: A Survey, Proc. IEEE, vol. 83, no. 5, p. 705-740.
- Kruger, J. et al. (2004). Obtaining the Eigenface Basis. Disponível em: <http://cnx.org/content/m12531/1.3/>. Acesso em 20 out. 2012.
- Sharmaa, A., Paliwala, K. (2012). A two-stage linear discriminant analysis for face-recognition. Pattern Recognition Letters. Vol. 33, Issue 9, 1 July 2012, p. 1157–1162
- Shakhnarovich, G. et al. (2004). Face Recognition in Subspaces. Handbook of Face Recognition, Eds. Stan Z. Li and Anil K. Jain, Springer-Verlag, Dezembro. 2004
- Sirovitch, L.; Kirby (1987), M. Low-dimensional procedure for the characterization of human faces. Journal of Optical Society of America, v. 4, p. 519-524.
- Turk, M.; Pentland, A (1991). Eigenfaces for Recognition. J. Cognitive Neuroscience, vol. 3, no. 1, 1991.
- Yale Univeristy (2012): Department of Computer Science. Center for Computational Vision and Control. Disponível em: <http://cvc.yale.edu/>. Acessado em 10/11/2012