

# SISTEMA DE RECONHECIMENTO DE FACES UTILIZANDO O MÉTODO DAS COMPONENTES PRINCIPAIS

*José Augusto Furtado Real\**  
*Thierry Pinheiro Moreira\*\**

## RESUMO

O assunto que trata do reconhecimento automático de faces humanas tem recebido grande atenção nos últimos anos, não somente devido ao desenvolvimento tecnológico, mas também pelo interesse comercial crescente neste setor. Porém, essa tarefa permanece um desafio para a área de reconhecimento de padrões. Esse desafio reside no fato de que as faces humanas são muito similares dependendo das condições de iluminação e pose. Este trabalho propõe um sistema automático de reconhecimento de faces baseado no método das componentes principais (PCA - Principais Componentes Analysis) robusto a variações que ocorrem na iluminação e orientação da face. O sistema combina um método de segmentação baseado no espaço de cor HSV (Hue-Saturation-Value) para extrair a região da face e colocá-la no centro da nova imagem em um tom de cinza conhecido. O método apresentou um alto desempenho no reconhecimento quando as faces estão sujeitas a mudanças de iluminação e variações elevadas na sua orientação.

**Palavras-Chave:** Reconhecimento de Faces. Detecção de Faces. Análise das Componentes Principais. Segmentação. Processamento de imagem.

## **FACE RECOGNITION SYSTEM USING PRINCIPAL COMPONENTES METHOD**

### **ABSTRACT**

*The issue of automatic recognition of human faces has received great attention recent years, not only because the technologic development, but also due to the growing commercial interest in the sector. However, this task stills a challenge for the pattern recognition area. This challenge resides in the fact that the human faces are very similar under variation of illumination and poses with large rotation angles. This work presents an automated face-recognition system based on principal components analysis (PCA) to be robust the variations that can occur in face illumination, and orientation. The system combines color image segmentation based on HSV (Hue-Saturation-Value)*

\* Universidade da Amazônia – UNAMA. Centro de Ciências Exatas e Tecnologia. e-mail: [real@unama.br](mailto:real@unama.br)

\*\* Universidade da Amazônia – Unama. Centro de Ciências Exatas e Tecnologia. e-mail: [thierry@unama.br](mailto:thierry@unama.br)

*color space to extract the face region on the face. The algorithms put the faces at the centre of the new image with a known grey tone. The system showed the efficient performance in face recognition under variation of illumination and poses with large rotation angles.*

**Keywords:** *Faces Recognition. Face Detection. Principal Component Analysis. Segmentation. Image processing.*

## 1 INTRODUÇÃO

A face possui um papel importante na comunicação entre os seres humanos. Através dela é possível não somente reconhecer quem é a pessoa, sua idade e a sua origem étnica, mas também o seu estado emocional. Todas estas características têm motivado a utilização da face em aplicações que envolvem segurança e geralmente associadas a outras formas de identificação biométrica, como o reconhecimento de voz e de impressões digitais. Além disso, em um futuro próximo, as máquinas poderão ser utilizadas para interagir com o ser humano, levando em consideração o seu estado emocional obtido através das expressões faciais (LI, 2005; HJELMÅS, 2001; YANG, 2002).

A área de pesquisa em reconhecimento de faces pode ser dividida em duas principais categorias: as baseadas em análise global (KARLEKAR, 1999) e as baseadas em características (análise local) (WANG, 2006). Nos métodos que utilizam uma forma de análise global, a imagem completa é transformada em um vetor e em seguida usada como padrão de entrada para algum método de reconhecimento. Nos métodos baseados em características, partes da imagem (ou pontos de controle) são extraídas para ser utilizadas como informações nos métodos de reconhecimento.

Em geral, independentemente da técnica utilizada, o processo de reconhecimento envolve quatro passos: detecção da face (localização), alinhamento (normalização), extração de características e classificação. Na detecção, o objetivo é verificar a existência de qualquer face ou elemento facial (boca, nariz etc.) na imagem, e caso exista, retornar a sua posição (MCKENNA, 1996; GARCIA, 1996). No alinhamento, as variações de luminância são minimizadas para reduzir sua influência na fase de reconhecimento da face (KOUZANI, 2005; LIN, 2005). Além disso, alguns métodos necessitam que as faces sejam centralizadas para manter uma coerência em relação ao banco de imagens, pois não são robustos a translação, como o que ocorre com os métodos estatísticos.

Na extração de características são eliminadas as redundâncias nos dados de tal forma que sejam preservadas somente as informações que são mais relevantes para a classificação. Através da extração de características é possível associar, de forma unívoca, cada face a um conjunto de informações que estão associadas a ela (WANG, 2006). O último passo do processo de reconhecimento é selecionar a face, do banco de imagens, que melhor se aproxima de uma imagem de teste (problema de reconhecimento de faces) ou comparar com padrões de expressões faciais (problema de reconhecimento de expressões faciais) ou rastrear uma determinada face em um ambiente com várias faces (GÓES, 2006).

O sistema proposto é um método de análise global baseado nas componentes principais (PCA – Principal Components Analysis). O primeiro trabalho que utilizou a técnica PCA para reconhecimento de faces foi proposto por Turk e Pentland (TURK, 1991a). Este método é hoje uma das técnicas mais utilizadas para reconhecimento de faces, também conhecido como “método baseado nas autofaces” (Eigenfaces method). Nesta técnica, cada imagem é convertida em um vetor e o conjunto de todos os vetores-imagens forma uma matriz (os vetores-imagem são organizados em linhas ou colunas). Em seguida é calculada a matriz de covariância das imagens, na qual são extraídas somente as componentes principais associadas aos maiores autovalores (maiores variâncias). As demais são descartadas devido influenciarem pouco no processo de classificação. As imagens projetadas nas componentes principais, denominadas de autofaces, são em seguida armazenadas em um banco de imagens. Para recuperar um padrão a partir de uma imagem de teste, esta imagem é projetada neste espaço e em seguida é utilizada uma métrica (ex: distância euclidiana, distância de Mahalanobis etc.) para determinar a menor distância entre esse padrão e os do banco de imagens.

No sistema proposto, as faces foram obtidas de um banco de imagens fornecido pelo Audio Visual Technologies Group (GTAV) (TARRÉS, 2010). As imagens desse banco são todas coloridas com resolução 240x320 e criadas no formato bitmap. O banco possui um total de 44 pessoas com 27 imagens por pessoa em diferentes poses, ângulos de visão e iluminação. As imagens de cada indivíduo é pré-processada para reduzir os efeitos da iluminação e centralizadas para minimizar o erro devido ao posicionamento em relação à câmera. Em seguida o método das componentes principais é aplicado em parte do banco (padrão de treinamento) para geração das autofaces (eigenfaces). As imagens selecionadas para teste são projetadas no autoespaço (eigenspace) e comparadas com as do banco para determinar a menor distância entre elas.

O trabalho foi organizado da seguinte forma: na seção 2 é apresentado o algoritmo de reconhecimento de faces baseado no método das componentes principais. Na seção 3 é descrito método de segmentação baseado no espaço de cores HSV. A seção 4 descreve o sistema proposto e na seção seguinte são apresentados os resultados experimentais do sistema com o banco de imagens da GTAV. A última seção é utilizada para traçar algumas conclusões em relação à capacidade do sistema a partir dos resultados obtidos na seção anterior.

## 2 MÉTODO DAS COMPONENTES PRINCIPAIS NO RECONHECIMENTO DE FACES

Considere uma imagem  $\chi$ , representada na escala de cinza, de dimensão  $N \times N$ , onde cada elemento corresponde à intensidade de um pixel da imagem. Para aplicação do método, cada uma das matrizes imagens é transformada em um vetor coluna de dimensão  $N^2 \times 1$ , representada por  $U_i$  onde  $i = 1, \dots, n$  e  $n$  é o número de imagens de treinamento (TURK, 1991b; PHIASAI, 2001).

Deseja-se representar cada imagem em um espaço de dimensão reduzida, da forma:

$$U_i = w_1\varphi_1 + w_2\varphi_2 + \dots + w_k\varphi_k$$

onde  $k \ll N^2$  e  $i = 1, \dots, n$ . Para reduzir o ruído, calcula-se o vetor média das imagens, da forma:

$$\mu = \frac{1}{n} \sum_{j=1}^n U_j$$

Subtraindo cada imagem da média, obtém-se:

$$\Delta_i = U_i - \mu$$

Observe que cada linha da matriz formada pelos vetores coluna  $\Delta_i$  corresponde aos elementos da imagem de uma posição em particular e as colunas as imagens de treinamento.

A covariância da matriz de vetores de imagem é calculada da forma:

$$C_{\Delta} = \frac{1}{n} \sum_{j=1}^n \Delta_j \Delta_j^T = \Omega \Omega^T$$

Os autovetores da matriz de covariância com maiores autovalores capturam a maior variação entre os vetores de treinamento  $\Delta_i$  e os de menores autovalores capturam as menores variações. A dimensão do espaço de entrada pode ser reduzida, criando-se um subespaço linear formado pelos autovetores com os maiores autovalores definida como componentes principais. Portanto, para a determinação das componentes principais é necessário calcular os autovalores e autovetores da matriz  $\Omega \Omega^T$ . Entretanto, a dimensão da matriz  $\Omega \Omega^T$  é  $N^2 \times N^2$  o que aumenta significativamente o tempo de cálculo dos autovalores, pois existem  $N^2$  autovalores. Uma solução alternativa é calcular os autovalores da matriz  $\Omega^T \Omega$  cuja dimensão é  $k \times k$ , que é muito menor que  $N^2$  e possui autovalores. Os autovalores da matriz correspondem aos maiores autovalores da matriz como mostrada em Turk, 1991(a) e Turk, 1991(b).

Considere que  $k$  autovetores normalizados ( $\|\varphi_i\| = 1$ ), denotado por  $\varphi_1, \varphi_2, \dots, \varphi_k$ , sejam selecionados correspondendo aos maiores autovalores, denominados de autofaces. Cada imagem de face do conjunto de treinamento pode agora ser representada nesta base, da forma:

$$\Delta_i = w_{1i} \varphi_1 + w_{2i} \varphi_2 + \dots + w_{ki} \varphi_k = U_i - \mu$$

onde  $w_{ji} = \varphi_j^T \Delta_i$  para  $i = 1, \dots, n$  e  $j = 1, \dots, k$ , corresponde a representação do vetor imagem  $\Delta_i$  na base  $\varphi_1, \varphi_2, \dots, \varphi_k$ .

O reconhecimento de uma determinada face de um conjunto de teste, denotada pelo vetor  $x$ , pode ser realizado calculando-se a representação deste vetor (subtraído da média  $\mu$ ), na base  $\varphi_1, \varphi_2, \dots, \varphi_k$ . Portanto,

onde  $\alpha_1, \alpha_2, \dots, \alpha_k$  são os coeficientes desta representação. A face que melhor aproxima é obtida calculando-se a distância da representação do vetor  $x$  para cada uma das representações das faces do conjunto de treinamento. Portanto,

$$d_i^2 = \sum_{j=1}^k (\alpha_j - w_{ji})^2$$

onde  $k$  é o número de componentes principais e  $i$  é o número de imagens de treinamento. Uma outra métrica utilizada é a distância de Mahalanobis, dada por:

$$d_i^2 = \sum_{j=1}^k \frac{1}{\lambda_j} (\alpha_j - w_{ji})^2$$

onde  $\lambda_j$  são os autovalores associados a cada um dos autovetores (autofaces).

### 3 ESPAÇO DE CORES HSV

Uma forma de detecção de faces é através da cor da pele devido à simplicidade no processamento e a sua robustez em relação às mudanças na geometria dos padrões de face (CAETANO, 2001; HSU, 2002; SANDEEP, 2002). Além disso, a cor da pele, mesmo considerando as diferenças raciais, apresenta representações similares em um determinado padrão de cor e dependendo da escolha do padrão, o desempenho do algoritmo de detecção pode ser afetado significativamente (KARLEKAR, 2000).

Existem vários métodos de especificação, criação e visualização de cores. Uma cor é geralmente especificada utilizando-se três coordenadas ou parâmetros que formam um determinado espaço de cores. Existem vários espaços de cores padrões, e a escolha por um deles depende de cada aplicação. Neste trabalho, foi utilizado o padrão de cor HSV (Hue, Saturation, Value) para detecção e centralização das faces. Este padrão baseia-se no matiz-saturação que o torna simples para especificar as propriedades das cores numericamente. Além disso, este padrão é invariante a mudanças de intensidade e orientação da fonte de luz. A conversão de uma imagem no padrão RGB (Red, Green, Blue) para o padrão HSV é calculada da forma:

$$H = \arccos \frac{\frac{1}{2}((R-G) + (R-B))}{\sqrt{(R-G)^2 + (R-B)(G-B)}}$$

$$S = 1 - 3 \frac{\min(R, G, B)}{R + G + B}$$

$$V = \frac{1}{3}(R + G + B)$$

onde R, G e B são cada uma das componentes do sistema RGB.

#### 4 DESCRIÇÃO DO SISTEMA

Um exemplo das imagens utilizadas para testar o sistema é mostrado na figura 1, observe que as faces são apresentadas em ângulos de  $0^\circ$ ,  $\pm 30^\circ$ ,  $\pm 45^\circ$ ,  $\pm 60^\circ$  e  $90^\circ$  em relação à câmera sobre um fundo sujeito a efeitos de sombra. Além disso, a face também pode estar deslocada na horizontal em relação ao centro da imagem, sofrer pequenas variações de escala (distância em relação à câmera) e também de iluminação.



**Figura 1** – Imagem em diferentes ângulos, iluminação e sombra.

O primeiro passo no processo de reconhecimento pelo sistema proposto é segmentar a face utilizando o formato HSV. Decompondo os planos de cores da imagem nas suas componentes H, S e V, e considerando somente a componente S, como mostrado no exemplo da figura 2, é possível facilmente extrair a face do fundo da imagem.



**Figura 2** – Componente S do padrão HSV.

Após a extração da face, uma nova imagem é criada em tons de cinza. Esta imagem é centralizada com o fundo assumindo sempre o nível mínimo de intensidade na escala de cinza (preto absoluto). A figura 3 mostra a imagem já processada.



**Figura 3** – Imagem centralizada em tons de cinza.

As imagens processadas são utilizadas como padrão de teste para a aplicação do algoritmo baseado na análise das componentes principais. Observe que o fundo de todas as imagens é representado com um mesmo valor de intensidade (zero na escala de cinza), fazendo com que a face seja a informação mais importante para classificação.

Na fase de reconhecimento de um padrão, uma imagem selecionada no banco de imagens de teste é pré-processada (como mostrada anteriormente) e comparada com as imagens de treinamento. A imagem que resultar em menor erro será a candidata a face selecionada.

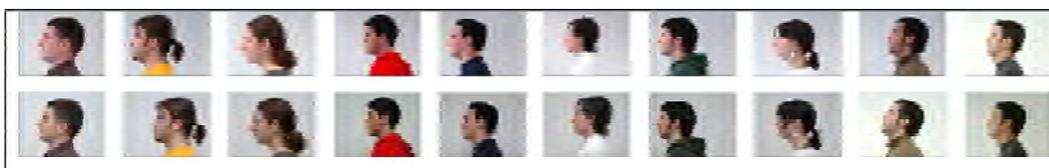
## **5 RESULTADOS EXPERIMENTAIS**

O sistema proposto foi avaliado utilizando o conjunto de imagens de treinamento da Audio Visual Technologies Group (TARRÉS, 2010) e autorizado para fins de pesquisa acadêmica. Este banco foi criado com o propósito de testar a robustez de algoritmos de reconhecimento de faces quando sujeito a mudanças na sua posição angular e condições de iluminação. O banco de dados possui um total de 44 pessoas com 27 imagens por pessoa, com variações no ângulo da face em relação à câmera e sujeitas a três tipos de iluminação (luz natural, luz intensa em 45° e luz com intensidade média em um ângulo quase frontal). A resolução das imagens é de 240x320 e estão no formato bitmap. Para atender aos objetivos do experimento, foram utilizados para treinamento e teste somente as imagens do banco que não apresentavam algum tipo de oclusão (face parcialmente coberta). A figura 4 mostra um exemplo de faces em posição frontal (0° em relação à câmera) pertencentes a este banco e sujeitas a variação na iluminação. A figura 5 mostra outro conjunto de imagens, também pertencentes a este banco, en-

tretanto posicionadas em um ângulo de 90° em relação à câmera. Uma parte das imagens do banco foi escolhida para treinamento, isto é, para determinação das componentes principais e outra parte utilizada para teste.



**Figura 4** – Amostra de 10 indivíduos pertencentes ao banco de treinamento e teste com variação de iluminação.



**Figura 5** – Amostra de 10 indivíduos pertencentes ao banco de treinamento e teste com variação de posição em relação ao centro e iluminação.

Os experimentos foram realizados para investigar os seguintes aspectos: a robustez do sistema em relação ao número de componentes principais, a robustez em relação à posição da face na imagem e às variações de iluminação.

Como é conhecido da teoria, método das componentes principais (MCP) é pouco robusto em relação à translação dos padrões. Quando o sistema foi aplicado com imagens sem qualquer tratamento prévio, resultou em um percentual de acerto próximo de 64%. Para obter uma alta taxa de acerto, as imagens do banco foram convertidas para o padrão HSV, segmentadas e transladadas para o centro, como mostrada na seção 4. Como a imagem do fundo não tem importância no processo de reconhecimento, foi substituída por uma tonalidade de cinza com valor conhecido. O próprio método das componentes principais irá descartar as informações do fundo após a execução. Com este pré-processamento, obteve-se um percentual de 94,94%.

As imagens que resultaram em insucesso são mostradas na figura 6. Observe que as duas primeiras colunas resultaram em acerto, se for considerado que a imagem recuperada corresponde à pessoa selecionada. Tendo em vista este aspecto, a taxa de acerto aumenta para 96,97%.



**Figura 6** – A primeira linha corresponde aos padrões selecionados e os da segunda os reconhecidos pelo sistema.

Foi investigado também o número mínimo de componentes principais no qual a taxa de acerto alcança o máximo. A tabela 1 mostra o percentual de acerto obtido quando o número de componentes principais varia. Observe que o número de componentes não tem influência após uma determinada quantidade. No caso do experimento, esta taxa ficou invariante acima de aproximadamente 40 componentes.

**Tabela 1** – Efeito do número de componentes na taxa de acerto.

<b>Número de Componentes Principais</b>	<b>Percentual de Erro</b>
<b>90</b>	<b>94,94</b>
<b>80</b>	<b>94,94</b>
<b>70</b>	<b>94,94</b>
<b>60</b>	<b>94,94</b>
<b>50</b>	<b>94,94</b>
<b>40</b>	<b>93,93</b>
<b>30</b>	<b>91,92</b>
<b>20</b>	<b>90,90</b>
<b>10</b>	<b>84,84</b>
<b>05</b>	<b>68,68</b>
<b>02</b>	<b>32,32</b>
<b>01</b>	<b>9,1</b>

Fonte: Resultado da Pesquisa

## 6 CONCLUSÃO

Neste artigo, foi desenvolvido um sistema baseado no método das componentes principais capaz de reconhecer faces sujeitas a variações de ângulo e de iluminação. As imagens no padrão HSV foram centralizadas e o tom de cinza do fundo da cena alterado de tal forma que apresentasse uma alta correlação entre eles (o fundo da imagem). Com esta estratégia observou-se que o percentual de acerto foi da ordem de 95%. Porém, se for considerado que alguns padrões, embora contabilizados como insucesso na classificação, correspondem às mesmas pessoas (em outros ângulos em relação à câmera) a taxa de acerto sobe para 97%. Foi também investigado o efeito do número de componentes principais na taxa de acerto. Observou-se que, para o banco de imagens utilizado, um valor acima de 40 componentes não resulta em melhoria na taxa de acerto, permitindo uma redução significativa do espaço de entrada.

## REFERÊNCIAS

CAETANO, Tibério; D. OLABARRIAGA, Silvia, A. C. BARONE, Dante. Performance Evaluation of Single and Multiple-Gaussian Models for Skin Color Modelling. IEEE INTERNATIONAL CONFERENCE ON IMAGE ANALYSIS AND PROCESSING, 11., Palermo, **Anais...**, Washington: IEEE Computer Society, 2001. p. 275-282.

GARCIA, Christophe; ZIKOS, G; TZIRITAS, George. **Face detection in color images using wavelet packet analysis**. 1. Ed. Washington: IEEE Computer Society, 1999. p. 703-708.

GÓES, Jucelaydi A.; SIQUEIRA, Caroline S. **Sistema de rastreamento de movimento urbano na Região Metropolitana de Belém**. 2006. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação) – Universidade da Amazônia, Belém, 2006.

HJELMÅS, Erik; KEE, Boon. **Face detection: a Survey**. Computer Vision and Image Understanding, West Lafayette, v. 83, n. 3, p. 236–274, 2001.

HSU, Rein-Lien; ABDEL-MOTTAZEB Mohamed; JAIN, Anil K. Face detect in color image. In: **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Ithaca, v. 24, n. 5, p. 961-706, ago. 2002.

KARLEKAR, J.; B. DESAI, U. Finding faces in color images using wavelet transform. In: INTERNATIONAL CONFERENCE ON IMAGE ANALYSIS AND PROCESSING, 10., 1999, Veneza, **Anais...**, Washington: IEEE Computer Society, 1999. p.1085-1088.

KARLEKAR, Jayashree; B. DESAI, Uday. Finding faces in wavelet domain for content-based coding of color images: two approaches. In: INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, 25., Istambul, **Anais...**, Istambul, 2000. p. 712-719.

KOUZANI, Abbas. Lighting normalisation for face recognition. In: IEEE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, 12, Gênova. **Anais...** Washington: IEEE Computer Society, 2005. p. 930-933.

LI, Stan Z.; JAIN, Anil K. **Handbook of Face Recognition**. 1. ed. Nova Iorque: Springer, 2005. 395 p.

LIN, Yuan-Pin et al. Webcam Mouse Using Face and Eye Tracking in Various Illumination Environments. In: ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY, 27., Shanghai, **Anais...**, Washington: IEEE Computer Society, 2005. p. 3738-3741.

MCKENNA, Stephen; GONG, Shaogang. Tracking Faces. In: INTERNATIONAL CONFERENCE ON AUTOMATIC FACE AND GESTURE RECOGNITION, 2., 1996, Killington, **Anais...**, Washington: IEEE Computer Society, 1996. p. 271.

PHIASAI, T.; ARUNRUNGRUSMI, S.; CHAMNONGTHAI, K. Face recognition system with PCA and moment invariant method. In: CIRCUITS AND SYSTEMS, 7., Sydney, **Anais...**, Washington: IEEE Computer Society, 2001. p. 165 -168.

SANDEEP, K.; N. RAJAGOPALAN, A. Human Face Detection in Cluttered Color Images Using Skin Color and Edge Information. In: INDIAN CONFERENCE IN COMPUTER VISION, GRAPHICS AND IMAGE PROCESSING, 3., 2002, Ahmedabad, **Anais...**, Ahmedabad.

TARRÉS, Francesc.; RAMA, Antonio. **GTAV Face Database**. Disponível em: <http://gps-tsc.upc.es/GTAV/ResearchAreas/UPCFaceDatabase/GTAVFaceDatabase.htm>. Acesso em: 3 jan. 2010.

TURK, M.; PENTLAND, A. Eigenfaces for recognition. **Journal of Cognitive Neuroscience**, Cambridge, v. 3, n. 1, p. 71-86, dez. 1991.

\_\_\_\_\_. Face recognition using eigenfaces. In: IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 3., Fort Collins, Washington: IEEE Computer Society, 1991. p. 586-591.

WANG Jing-Wein. Efficient Facial Component Extraction for **Detection and Recognition**. In: 18TH INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, 2. Hong Kong. **Anais...**, Washington: IEEE Computer Society, 2006. p. 1208-1211.

YANG, Ming-Hsuan; KRIEGMAN, David J.; AHUJA, Narendra. **Detecting faces in Images: a Survey**. IEEE Transactions on Pattern Analysis and Machine Intelligence, Ithaca, v. 24, n. 1