

# SISTEMA DE RECONHECIMENTO FACIAL BASEADO EM ANÁLISE DE COMPONENTES PRINCIPAIS

Ernesto Luiz Andrade Neto e Lee Luan Ling

DECOM-FEEC-UNICAMP

P.O. Box 6101

13081-970 - Campinas - SP - BRAZIL

fone : +55-19-788-2053 fax : +55-19-289-1395

{ernestoa, zhao}@decom.fee.unicamp.br

**Resumo** - Neste trabalho mostramos a implementação de um sistema de reconhecimento facial. O sistema de reconhecimento facial emprega a análise de componentes principais de um conjunto de imagens exemplo de faces para modelar a face. O modelo de componentes principais obtido na análise das imagens exemplo é utilizado para extrair características das imagens faciais a serem reconhecidas pelo sistema. A classificação das características obtidas é realizada com um classificador de distância mínima. O sistema obteve uma taxa de classificações corretas de 97,70% quando testado com 435 imagens de 102 indivíduos.

**Abstract** - We present the development and implementation of a face recognition system using principal component analysis to construct a face model from a training set of face images. The computed principal component face model is applied in the feature extraction task. We tested the system in the task of correctly classifying 435 images from 102 people with a minimum distance classifier. These images have a great deal of variations concerning position, scale, expression and illumination, even among images of the same person. The proposed system was able to handle such variations and achieved a recognition rate of 97.70%.

**Palavras Chaves** : reconhecimento facial, processamento de imagens, análise de componentes principais.

## 1. INTRODUÇÃO

A capacidade de reconhecer faces desempenha um papel muito importante em nossas vidas. Entretanto, traduzir nossa capacidade de reconhecer faces em um sistema automático de reconhecimento não é uma tarefa fácil, mesmo com o auxílio do computador. A face humana pode apresentar em uma imagem grandes variações em relação à parâmetros como posição, escala, iluminação ou expressão. Existem muitas aplicações para um sistema de reconhecimento facial que consiga aproximar a confiabilidade e a robustez da habilidade humana em reconhecer faces. Entre as várias aplicações dos sistemas de reconhecimento facial [1] podemos citar : controle de acesso, identificação de criminosos e tarefas gerais de segurança. Além das aplicações relacionadas principalmente com serviços de segurança as técnicas de reconhecimento facial podem ser empregadas em interfaces homem-máquina [2] [3], permitindo uma interação mais amigável entre usuários e computadores.

As tentativas pioneiras de reconhecer faces humanas em computadores baseavam-se em medidas de distâncias sobre pontos característicos da face humana [4]. Os trabalhos mais recentes desenvolvidos para o reconhecimento de faces analisam imagens faciais digitalizadas [1]. A comparação dos métodos de reconhecimento de face que utilizam padrões extraídos de imagens faciais com os que empregam medidas de características sobre a imagem facial mostra que os padrões extraídos na análise da imagem facial apresentam um melhor desempenho para o reconhecimento facial [5]. Atualmente existe uma grande variedade de técnicas empregadas para a análise e reconhecimento de imagens faciais. O método de análise de componentes principais de um conjunto de imagens faciais empregado em nosso sistema de reconhecimento facial foi descrito primeiramente em [6]. É um método baseado em análise das estatísticas da distribuição dos pixels de um conjunto de imagens faciais. O método de componentes principais aplicado a imagens faciais é desenvolvido com base nos trabalhos de Kirby e Sirovich [7] [8] sobre formas eficientes de comprimir e reconstruir imagens com a análise de componentes principais. A análise de componentes principais de imagens faciais apresenta uma relação com as técnicas de reconhecimento que empregam memórias autoassociativas [9]. Na tarefa de reconhecimento facial também são empregadas outras arquiteturas de redes neurais. Entre as mais eficientes podemos citar o casamento dinâmico de ligações [10] e as árvores de grupos de redes neurais [11].

Neste trabalho é descrito o modelo de representação da face, detalhando o emprego da análise de componentes principais de imagens faciais. Em seguida é abordada a implementação do sistema de reconhecimento facial. Na sequência são mostrados os resultados do sistema. Finalmente, são apresentados os comentários e conclusões sobre o desempenho do sistema de reconhecimento facial.

## 2. MODELO DE REPRESENTAÇÃO DA FACE

Na construção do modelo de representação da face com o método de análise de componentes principais devemos escolher a posição ou expressão da face na imagem que vai ser representada pelo modelo facial. Esta é uma maneira de tratar as grandes variações presentes na face, pois construímos um modelo que representa mais precisamente um determinado tipo de face. Mesmo com a restrição de modelar um determi-

nado tipo de face, o modelo facial escolhido permite grandes distorções para novas faces apresentadas. As faces escolhidas compõem um conjunto de imagens que será submetido a análise de componentes principais. A análise de componentes principais fornece os autovalores e autovetores do conjunto de imagens faciais analisadas. Os autovetores deste conjunto de imagens exemplo formam uma base para a projeção de novas imagens faciais. Cada imagem facial é codificada em termos de sua projeção sobre os autovetores obtidos na análise de componentes principais.

No desenvolvimento da análise de componentes principais do conjunto de imagens exemplo consideramos uma imagem facial representada por  $I(x, y)$  uma matriz bidimensional de níveis de cinza de dimensão  $a \times b$ . As imagens bidimensionais são transformadas para a forma de vetores linha de dimensão  $(a \times b) \times 1$ . O conjunto de  $M$  imagens de treino na forma de vetores linha é dada por

$$S = \{I_1, I_2, I_3, \dots, I_M\} \quad (1)$$

A imagem média do conjunto de treino é

$$\Psi = \frac{1}{M} \sum_{n=1}^M I_n \quad (2)$$

onde  $M$  é o número de imagens do conjunto de treino.

Definimos o vetor linha da imagem do conjunto de treino normalizada em relação a média do conjunto de treino como

$$\Phi_i = I_i - \Psi \quad i = 1, \dots, M \quad (3)$$

Com as imagens normalizadas em relação a média formamos a matriz  $A$ , definida como

$$A = [\Phi_1 \quad \Phi_2 \quad \dots \quad \Phi_M] \quad (4)$$

Obtemos então a matriz de autocorrelação do conjunto de imagens definida como

$$C = AA^T \quad (5)$$

Os autovalores,  $\lambda_i$ , e autovetores,  $u_i$ , da matriz  $C$ , satisfazem a seguinte relação:

$$Cu_i = \lambda_i u_i \quad (6)$$

Em princípio a matriz  $C$  possui  $a \times b$  autovetores o que dependendo das dimensões da imagem pode inviabilizar a análise. Podemos simplificar o cálculo dos autovetores e autovalores nos valendo do fato de que quando o número de imagens presentes,  $M$ , é muito menor que a dimensão dos vetores das imagens,  $a \times b$ , temos apenas  $M-1$  autovetores significativos os restantes estarão associados a autovalores nulos. Na análise deste novo modelo encontramos os autovalores,  $\mu_i$ , e os autovetores,  $v_i$ , da matriz  $A^T A$  tais que

$$A^T A v_i = \mu_i v_i \quad (7)$$

Pré-multiplicando ambos os membros de (7) por  $A$  obtemos

$$AA^T A v_i = \mu_i A v_i \quad (8)$$

Característica do Grupo	Número de Imagens
Modelo	193
Frontal	319
Rotacionada	116
Sorrindo	142
Óculos	62

TAB. 1: Distribuição da imagem da base de dados segundo as características presentes nas imagens

De (8) podemos notar que  $A v_i$  são os autovetores da matriz  $C = AA^T$  e  $\mu_i$  representa seus autovalores não nulos.

Então para obtermos os autovetores  $u_i$  de  $C$  realizamos combinações lineares de  $v_i$  e  $A$  tais que

$$u_i = \sum_{k=1}^M v_i \Phi_k \quad i = 1 \dots M \quad (9)$$

pois  $A v_i$  é o  $i$ -ésimo autovetor  $u_i$  da matriz  $C$ . Desta forma reduzimos o tempo de cálculo dos autovetores significativos do conjunto de imagens faciais podemos então tratar imagens de grandes dimensões se  $M \ll a \times b$ .

Uma vez determinadas as componentes principais do conjunto de imagens exemplo a caracterização de uma nova face apresentada ao sistema, segundo o modelo definido pela componentes principais, é dada por

$$w_k = u_k^T (\Gamma - \Psi) \quad k = 1 \dots M' \quad (10)$$

onde  $w_k$  é o valor da projeção da nova imagem  $\Gamma$  subtraída da média do conjunto de imagens exemplo  $\Psi$  sobre o  $k$ -ésimo autovalor  $u_k^T$  e  $M'$  é a quantidade de autovalores em que a imagem vai ser projetada ( $M' \leq M$ ). A imagem codificada com o modelo de componentes principais pode então ser caracterizada como o conjunto de pesos de projeção sobre os autovetores do modelo definido via

$$\Omega^T = [w_1, w_2, \dots, w_M] \quad (11)$$

Representamos também as várias classes presentes na base de dados por outro conjunto de vetores de pesos de projeção sobre os autovetores do modelo de componentes principais definido como

$$\Omega_l^T = [w_1, w_2, \dots, w_M] \quad l = 1 \dots K \quad (12)$$

onde  $K$  é o número de classes na base de dados. Os elementos do vetor de pesos em (12) podem ser obtidos de uma imagem exemplo para cada indivíduo ou da média aritmética dos pesos de várias imagens de um indivíduo. A distância entre os vetores de características de uma classe e de uma nova uma imagem apresentada ao sistema é definida como

$$\varepsilon_i = \|(\Omega - \Omega_i)\|^2 \quad (13)$$

Para classificarmos uma nova imagem em uma das classes pré-definidas utilizamos um classificador de distância mínima que minimiza a distância entre o vetor de pesos de uma nova imagem  $\Omega$  e os vetores de pesos das classes pré-definidas  $\Omega_i$  (13). A classe pré-definida que apresenta a menor distância em relação a nova imagem é atribuída a nova imagem.



Figura 1: Exemplos de faces usadas como modelo para a análise de componentes principais.

### 3. IMPLEMENTAÇÃO DO SISTEMA DE RECONHECIMENTO

A construção de modelo facial baseado na análise de componentes principais foi realizada com um conjunto de imagens faciais frontais, sem óculos e com expressão neutra. Este tipo de imagem foi escolhida por usualmente aparecer em documentos de identificação. Na Figura 1 podemos observar alguns exemplos das imagens usadas no modelo facial do sistema de reconhecimento.

Podemos empregar o modelo facial obtido com as faces frontais para codificar e reconhecer novas faces, apresentadas ao sistema de reconhecimento. As novas imagens faciais podem possuir algumas variações em relação as imagens usadas

como modelo, não há portanto a necessidade de recompor-mos o modelo facial para pequenas variações de expressão ou posição. Para os testes do sistema de reconhecimento facial foi adquirido um conjunto de 435 imagens faciais de 102 indivíduos. Não foram feitas restrições severas quanto a posição, iluminação, expressões ou ao uso de óculos por parte dos indivíduos da base de dados durante a aquisição das imagens. Garantimos apenas que para cada indivíduo exista ao menos uma imagem frontal, sem óculos e com expressão neutra para a composição do modelo facial. As imagens adquiridas foram classificadas segundo as diversas características presentes nas imagens que são consideradas relevantes para a determinação do grau de tolerância à variação do modelo de componentes principais das imagens.

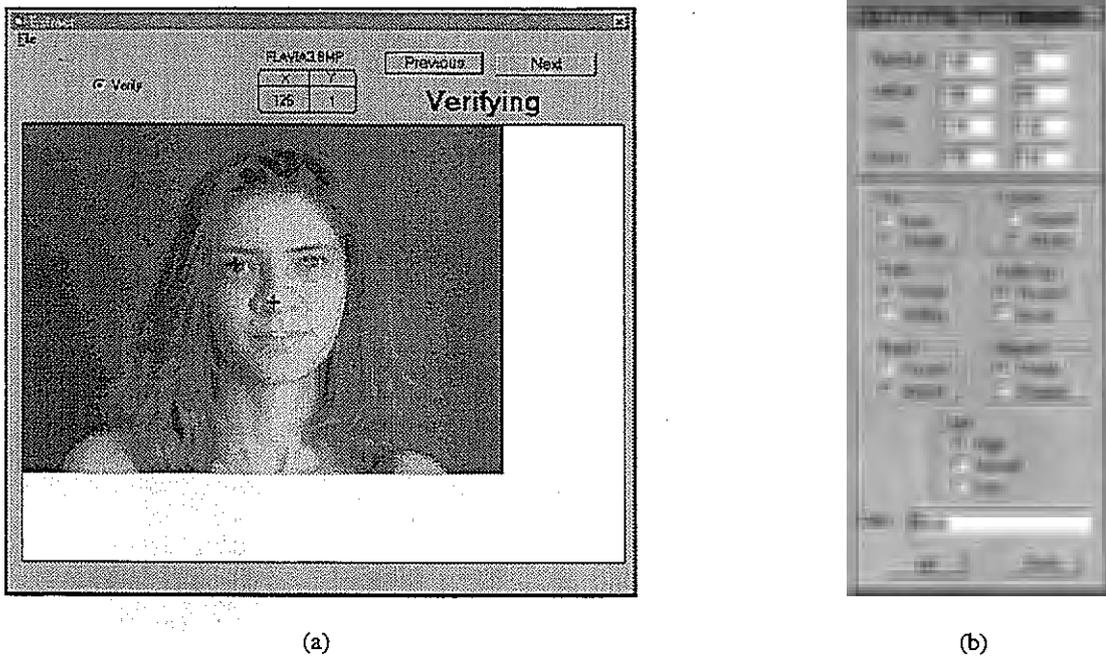


Figura 2: Software utilizado para registrar a posição dos olhos, nariz e boca (a) e as outras características da imagem facial como iluminação, expressão e posição (b).

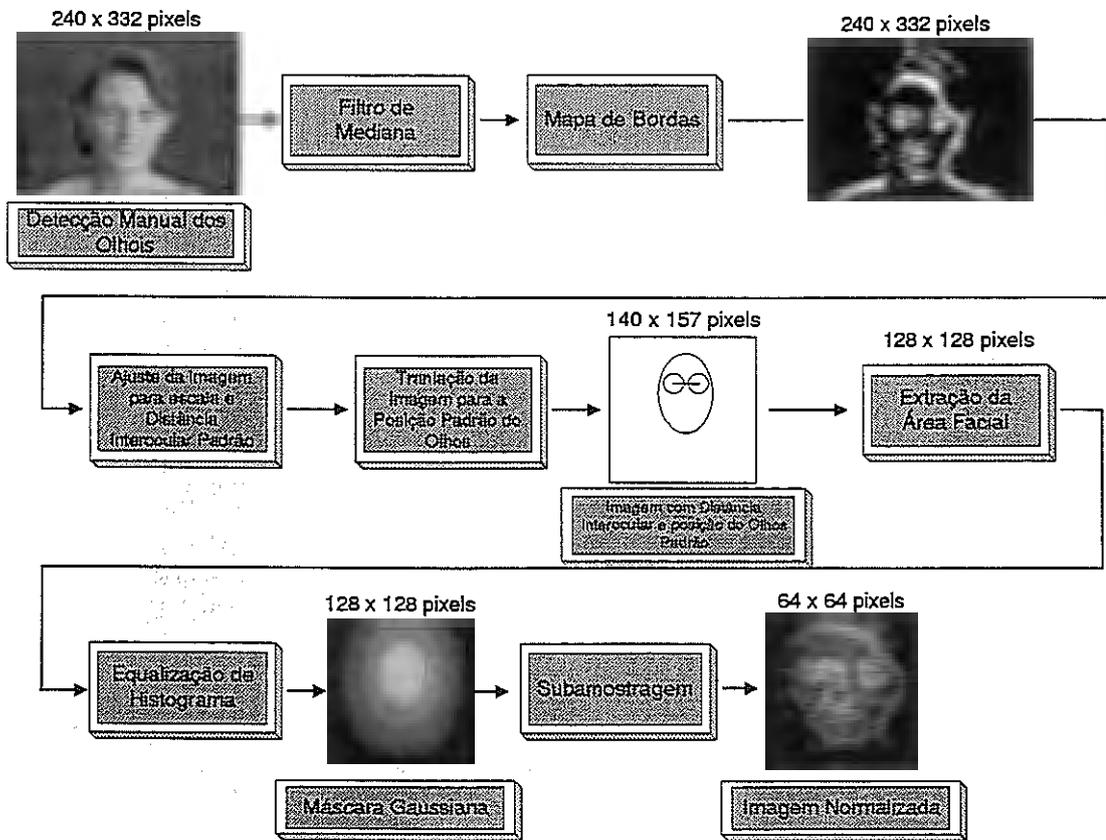


Figura 3: Pré-processamento e normalização das imagens faciais para extração de características e reconhecimento.

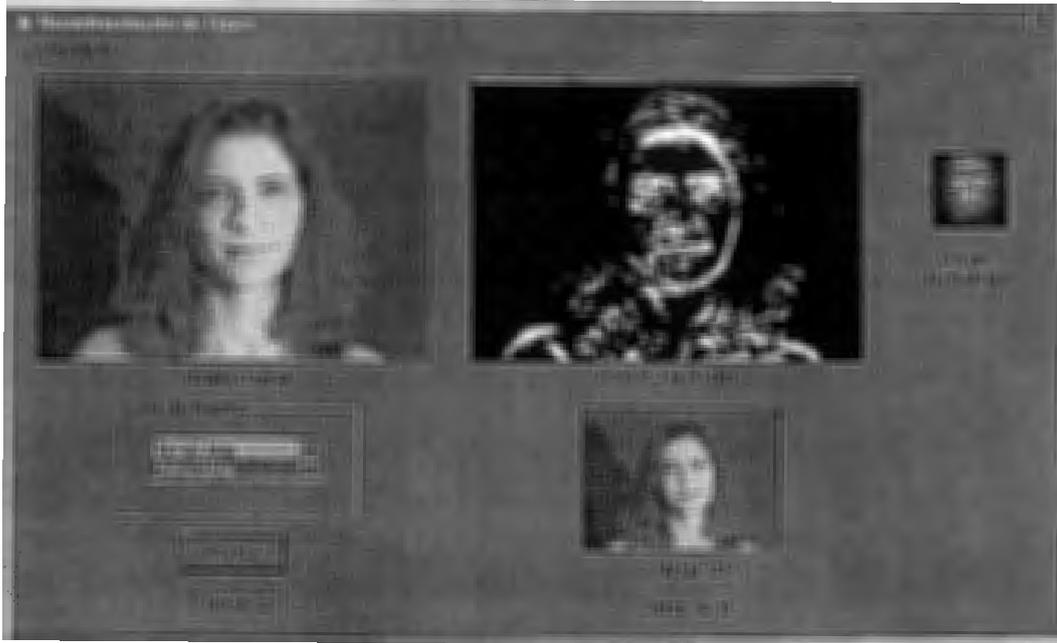
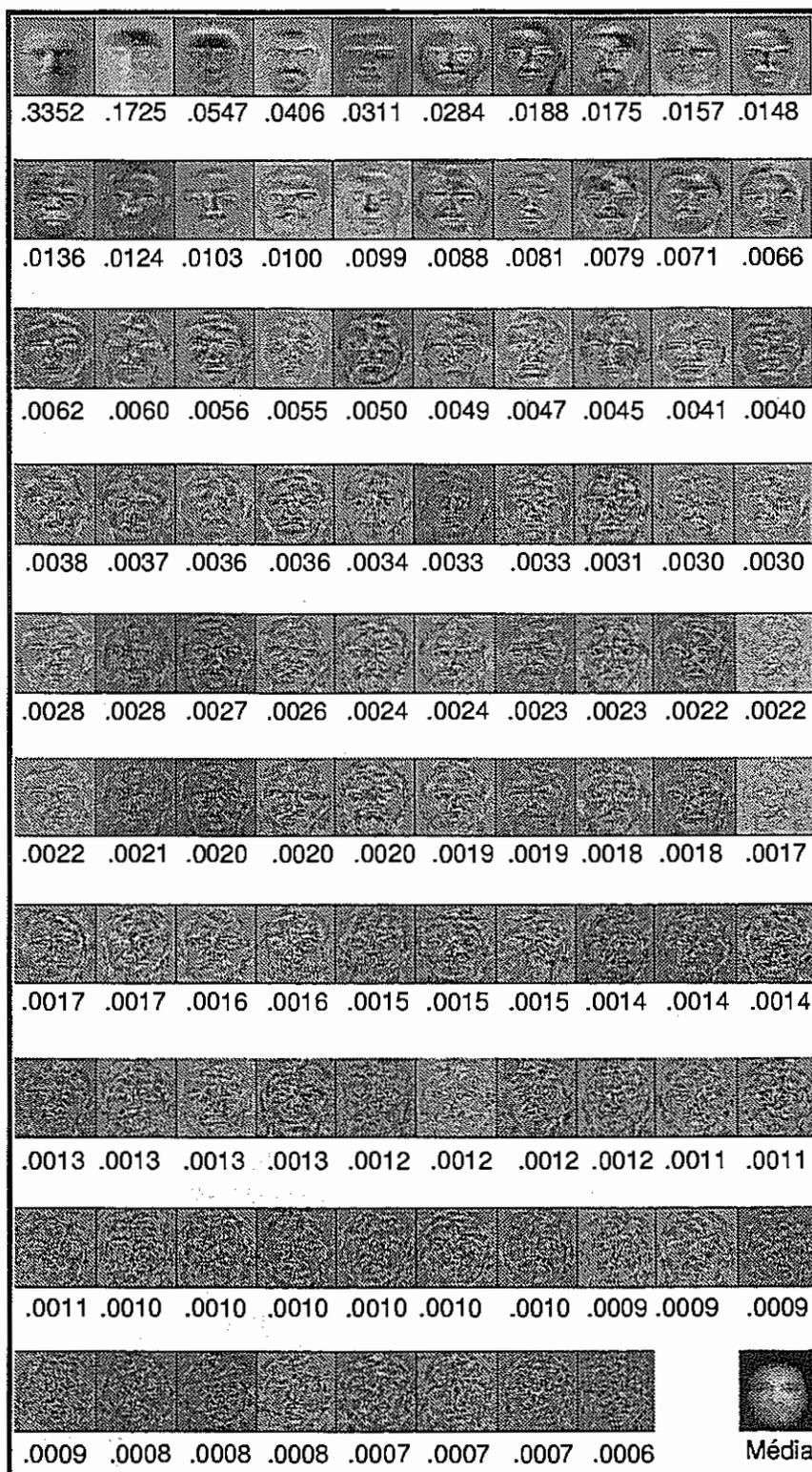


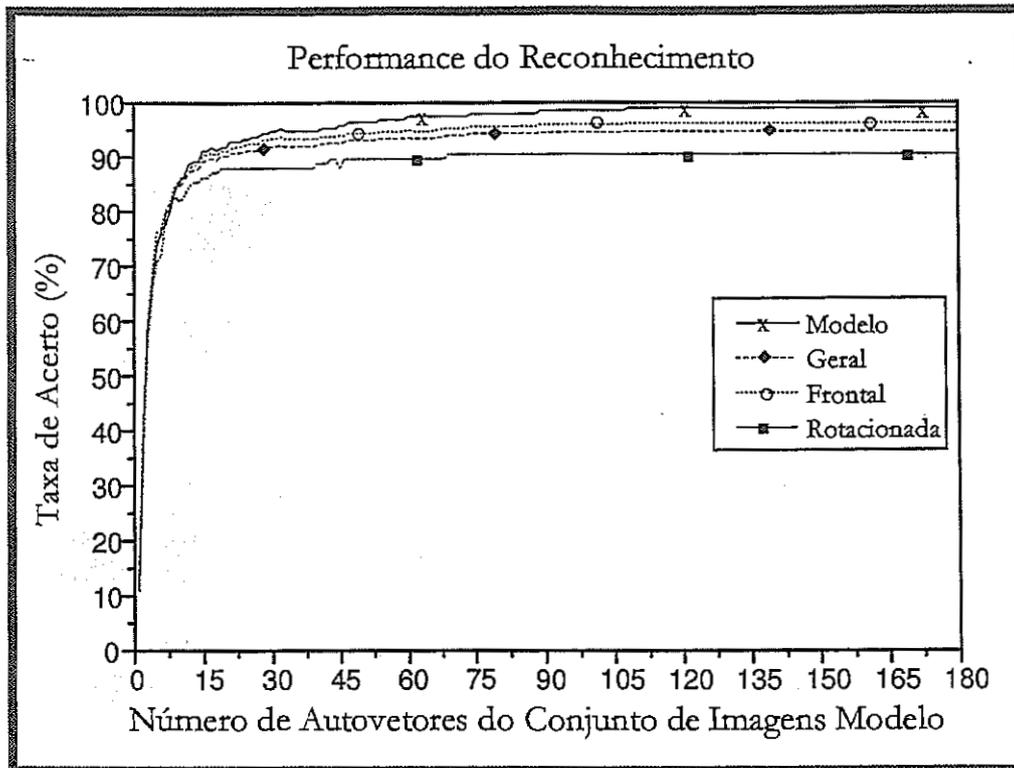
Figura 4: Tela do aplicativo de demonstração do sistema de reconhecimento facial.

Na Tabela 1 é mostrada a distribuição das imagens segundo estas características. O grupo modelo é formado por faces que são usadas para compor o modelo de componentes principais. O grupo frontal é composto por faces em posição frontal. O grupo rotacionada é composto por faces que apresentam uma rotação em relação a posição frontal da face. Os grupos óculos e sorrindo são compostos por imagens com estas características. Na Figura 2 observamos uma exemplo da utilização do aplicativo para o registro da posição dos olhos e catalogação da base de dados de imagens faciais. Antes de obtermos o modelo facial todas as imagens que são empregadas no sistema de reconhecimento facial devem ser pré-processadas e normalizadas. Podemos observar na Figura 3 o diagrama de blocos do tratamento dado às imagens faciais antes de serem analisadas pelo sistema de reconhecimento. No pré-processamento o filtro de mediana é usado para eliminar os ruídos no processo de aquisição da imagem. O mapa de bordas da imagem realça as características dos contornos da face e dos olhos. Partindo da posição dos olhos determinada manualmente na imagem original normalizamos a imagem pré-processada para uma posição e distância interocular padrão. Extraímos a área que contém a face na imagem normalizada e realizamos a equalização de histograma para compensar os efeitos da variação de iluminação entre as imagens. A máscara gaussiana aplicada a imagem facial é usada para diminuir a contribuição da informação do fundo da imagem e da área dos cabelos enfatizando a região dos olhos. Os olhos foram determinados como sendo um dos componentes mais significativos no reconhecimento de face pelos seres humanos [1]. O último passo de processamento é a redução da imagem por subamostragem dos pixels para podermos aumentar a velocidade de processamento do sistema de reconhecimento. Na Figura 4 é apresentada uma tela de demonstração do sistema de reconhecimento facial, onde se observa a escolha de uma

determinada face a ser apresentada ao sistema, os passos intermediários no pré-processamento e o resultado apresentado pelo sistema de reconhecimento facial. Em nosso modelo de imagens frontais, sem óculos e com expressão neutra realizamos a análise de componentes principais de 179 imagens exemplo. O resultado da análise pode ser visto na figura 5 onde temos os autovetores do conjunto de imagens exemplo os quais possuem a forma de faces. Os autovetores que possuem as formas mais definidas na Figura 5 detêm uma proporção maior da variância das imagens exemplo e representam características comuns a todo o conjunto de imagens faciais. Os autovetores que apresentam uma aparência mais ruidosa na Figura 5 detêm uma proporção menor da variância das imagens exemplo, representam características que diferenciam as imagens exemplo. Podemos interpretar as características obtidas na Figura 5 como uma representação das variações presentes na imagens exemplo. A informação extraída das variações nas imagens exemplo é utilizada para codificar e caracterizar novas imagens faciais. Nosso sistema de reconhecimento facial é composto pelos seguintes estágios: aquisição da imagem, realizada com câmera de vídeo e placa digitalizadora; pré-processamento e normalização da imagem; extração de características, utilizando o modelo de componentes principais obtido anteriormente das imagens exemplo selecionadas; e classificação, empregando o classificador de distância mínima. Todas as rotinas que compõem o sistema de reconhecimento facial foram implementadas em linguagem C++. O tempo de processamento total para uma imagem ser reconhecida pelo sistema é de cerca de 4s em um microcomputador PENTIUM™ de 100MHz. Sendo que não estão computados neste tempo a aquisição da imagem e a determinação manual da posição dos olhos na imagem para a normalização que é realizada em separado por um aplicativo específico para esta tarefa desenvolvido em Visual Basic™.



**Figura 5:** Componentes principais do conjunto de imagens de treino. O número sobre a imagem de cada autovetor representa a percentagem da variância normalizada da base de dados descrita pelo autovetor. A imagem média do conjunto está no canto inferior direito da figura.



**Figura 6:** Resultados do sistema de reconhecimento facial em função do número de características empregadas.



**Figura 7:** Erros de classificação do sistema de reconhecimento facial. O erro da face considerada como modelo está destacado.

#### 4. RESULTADOS

Nos testes do sistema de reconhecimento facial utilizamos 435 imagens dos 102 indivíduos presentes na base de dados adquirida. O centro de classe, que representa cada indivíduo no sistema de reconhecimento facial, empregado é a média aritmética das características de todas as imagens do indivíduo presentes na base de dados. Durante a classificação analisamos as distâncias entre os vetores de centro de classe pré-definidos e os vetores de características das novas imagens faciais. Para observarmos a tolerância à variações na imagem de nosso modelo de componentes principais, construído com imagens frontais, sem óculos e com expressão neutra, comparamos o desempenho de classificação do sistema de reconhecimento facial com imagens frontais, rotacionadas, imagens

usadas no modelo de representação e contra todas as imagens da base de dados independente de suas características. Os resultados destas comparações podem ser observados na Figura 6 onde observamos a variação da taxa de acerto de classificação para um variado número de autovetores utilizados como extratores de características. Na classificação das imagens que se assemelham ao modelo (193 imagens) obtivemos os menores erros, apenas 1,03%, estas imagens são bem representadas pelo sistema, pois compõem seu modelo facial. Na imagens frontais (319 imagens) o erro também foi pequeno 1,56%, neste caso o sistema de reconhecimento facial apresentou uma grande tolerância as variações já que haviam imagens com expressão e utilizando óculos no grupo frontal. O grupo rotacionada (116 imagens) é o que apresenta maior erro 4,31% pois este grupo é o que contém as imagens que mais se desviam do modelo das imagens exemplo. No geral (435 imagens) o sistema de reconhecimento facial apresentou um erro de apenas 2,29%. Os erros de classificação realizados pelo sistema de reconhecimento facial podem ser vistos na Figura 7. Notamos que 6 das 10 imagens que apresentaram erros estão utilizando óculos. Duas imagens com erros possuem uma grande rotação da face na imagem em relação a posição frontal. Os dois erros com imagens modelo são devidos principalmente a distorção no foco da imagem durante o processo de aquisição.

## 5. CONCLUSÕES

Mostramos neste trabalho um método eficiente de modelar imagens faciais que apresenta uma grande tolerância em relação a variações entre o modelo e as imagens representadas com este modelo. O processo de pré-processamento e normalização desenvolvido contribui para um aumento da performance do método de análise de componentes principais pois acentua as diferenças entre as faces de diferentes indivíduos. As taxas de acerto de classificação de nosso sistema de reconhecimento facial se encontram entre as taxas de acerto para as implementações de sistemas de reconhecimento facial descritas na literatura [10] [11] [12]. Devemos ressaltar o fato de que nosso sistema trabalha com uma base de dados que quando comparada aos outros sistemas de reconhecimento de face descritos possui poucos exemplos de cada indivíduo ou não tem restrições quanto as condições de aquisição das imagens faciais.

## REFERÊNCIAS

- [1] R. Chellappa, C. L. Wilson, S. Sirohey, Human and Machine Recognition of Faces : A Survey, Proceedings of the IEEE, Vol. 83, No 5, may, 1995.
- [2] H. Sako, M. Whitehouse, A. Smith e A. Sutherland, Real-Time Facial Feature Tracking Based on Matching Techniques and Its Applications, Proceedings of 12o IEEE ICPAR, Israel, pp 320-324, 1994.
- [3] O. Hasegawa, K. Yokosawa, M. Ishizuka, Real-Time Parallel and Cooperative Recognition of Facial Images for an Interactive Visual Human Interface, Proceedings of 12o IEEE ICPAR, Israel, pp 384-387, 1994.
- [4] A. Samal e P. A. Iyengar, Automatic Recognition and Analysis of Human Faces and Facial Expressions : A Survey, Pattern Recognition, Vol. 25, No 1, pp 65-77, 1992.
- [5] R. Brunelli e T. Poggio, Face Recognition : Features versus Templates, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, No 10, pp 1042-1052, october, 1993.
- [6] M. Turk, A. Pentland, Eigenfaces for Recognition, Journal of Cognitive Neuroscience, Vol. 3, No 1, pp. 71-86, 1991.
- [7] M. Kirby, L. Sirovich, Application of the Karhunen-Loève Procedure for the Characterization of Human Faces, IEEE Transactions Pattern Analysis and Machine Intelligence, vol. 12, pp 103-108, january, 1990.
- [8] L. Sirovich, M. Kirby, Low-Dimensional Procedure for the Characterization of Human Faces, Journal of Optical Society of America, Vol. 4, No. 3, pp 519-524, march, 1987.
- [9] H. Abdi, D. Valentin, A. J. O'Toole, A generalized Autoassociative Model for Face Processing and Sex Categorization : From Principal Component Analysis to Multivariate Analysis, Optimality in Biological and Artificial Networks, D. Levine (Ed.) in press, 1996.
- [10] R. P. Würtz, Building Visual Correspondence Maps - From Neuronal Dynamics to a face Recognition System, in Roberto Moreno-Díaz and José Mira-Mira (eds.), Brain Process, Theories and Models, pp. 420-429, MIT PRESS, november, 1995.
- [11] M. Zhang, J. Fulcher, Face Recognition Using Artificial Network Group-Based Adaptive Tolerance (GAT) Tress, IEEE Transactions on Neural Networks, Vol. 7, No. 3, pp. 555-567, may, 1996.
- [12] B. Moghaddam, A. Pentland, An Automatic System for Model-Based Coding of Faces, IEEE Data Compression Conference, Snowbird, Utah, march, 1995.

**Ernesto Luiz Andrade Neto** formou-se em Engenharia Elétrica pela UFSM em 1995, obteve o título de mestre em Ciência pela UNICAMP em 1997, ambos em Engenharia Elétrica. Atualmente está no programa de doutorado da Faculdade de Engenharia Elétrica e de Computação da UNICAMP. Seus interesses de pesquisa incluem Arquiteturas para a Simulação de Redes ATM, Reconhecimento Padrões e Redes de Comunicações (<http://www.mc21.fee.unicamp.br/ernesto/>).  
**Lee Luan Ling** formou-se em Engenharia Elétrica pela USP

em 1980, obteve o título de mestre em Ciência pela UNICAMP em 1984, o título de PhD pela Cornell University, EUA, em 1991 e o título de Professor Livre Docente pela UNICAMP em 1997, todos em Engenharia Elétrica. Desde 1984 é Professor da Faculdade de Engenharia Elétrica e de Computação da UNICAMP. Atualmente ocupa a posição de Professor Colaborador MS-4. Seus interesses de pesquisa incluem Reconhecimento Padrões e Redes de Comunicações (<http://www.lrprc.fee.unicamp.br/>).