

LOCALIZADOR DE PADRÕES FACIAIS: UMA APLICAÇÃO USANDO-SE REDES NEURAIS ARTIFICIAIS

*Adriano Martins Moutinho*¹
*Antonio Carlos Gay Thomé*²
*Pedro Henrique Coelho*³
*Luiz Biondi Neto*³
*Lidia Angulo Meza*⁴

Resumo: Um dos problemas comumente encontrados em sistemas de segurança, baseados em identificação facial, é o de separar a região facial propriamente dita e caracterizada por conter olhos, nariz e boca, do fundo e de outros objetos presentes na imagem. Localizar uma área da imagem correspondente a uma face é parte integrante de qualquer sistema automático de reconhecimento de identidade e constitui um problema bastante complexo podendo, inclusive, ser estudado separadamente. Este trabalho sugere um método, usando Redes Neurais Artificiais (RNA), que visa localizar e extrair de uma imagem completa, a região que contenha a face propriamente dita, em outras palavras localizar a face. Com o objetivo de encontrar a posição correta da face, é proposto um método de segmentação que subdivide a imagem em várias partes. São aplicadas também diversas técnicas de pré-processamento envolvendo correção de histograma e análise de componentes principais (PCA). A pesquisa também engloba o desenvolvimento de um método para correção de iluminação que, em última análise, melhora a capacidade de reconhecimento da rede neural. Finalmente, é apresentado um estudo de casos, envolvendo uma amostra de cerca de 5000 imagens, contendo faces e não faces, usadas para treinamento, e 400 imagens, diferentes das usadas no treinamento, para teste da arquitetura investigada, na qual é verificada a eficiência das técnicas de empregadas.

Palavras-Chave: Localizador de padrões faciais, Redes neurais artificiais, Processamento de imagens, Análise de componentes principais.

Abstract: Security systems based on face recognition often have a problem to separate the face region, containing nose, mouth and eyes, from the rest of the objects inside the image. Finding the right position of a face is part of any automatic identity recognizing system, and it is, by itself, a problem that is very difficult to solve, normally being handled separately. This paper offers a solution, using artificial neural networks (ANN), to extract the face position from a complete image. In order to accomplish this goal, a segmentation method was created and several image pre-processing steps, from histogram equalization to lightning correction, were used to improve neural network recognizing capability. Finally, 400 images were used as a "case study" to test the system and perform localization of some facial features like mouth, nose and eyes, besides the localization of the face itself.

Key words: Facial Pattern Locator, Artificial Neural Networks, Image Processing, Principal Components Analysis.

¹ IM / NCE - UFRJ

² Instituto de Matemática / DCC - UFRJ

³ DETEL/UERJ

⁴ UVA-RJ

1. INTRODUÇÃO

Neste trabalho define-se como face o menor retângulo que engloba os olhos, nariz e boca de uma imagem de um indivíduo hipotético, separando-os dos demais objetos que compõem a imagem. A figura 1, de forma geral, mostra um exemplo de face.



Figura 1 - Exemplo de face.

Mesmo na imagem mostrada na figura 1, que é bastante simples, identificar corretamente a posição da face pode ser uma tarefa bastante complexa. Muitas dificuldades podem surgir, tais como: a presença de óculos, bigode, barba, variações no tom de pele, rotações, diferentes inclinações da cabeça, etc.

Tornar, tanto quanto possível, o sistema de localização imune a estas variações é um dos objetivos deste trabalho.

Como a identificação de padrões faciais é um problema de classificação envolvendo dados altamente não lineares, torna-se adequada a solução usando redes neurais MLPs (*Multilayer Perceptron*) [1][4]. Assim, uma rede neural é treinada usando como entrada diversas seqüências contendo faces, de acordo com a defini-

ção preliminar mostrada na figura 1, bem como diversas outras imagens não contendo faces e constituídas basicamente de ruídos e de formas geométricas as mais variadas possíveis.

As figuras 2 e 3 mostram exemplos de bancos de dados de imagens contendo faces e não faces, respectivamente.

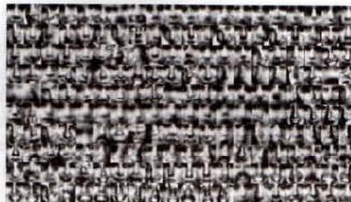


Figura 2 - Banco de dados de faces.



Figura 3 - Banco de dados de não faces.

2. VISÃO GERAL DO SISTEMA LOCALIZADOR DE FACES

O sistema de localização proposto pode ser dividido em quatro módulos principais mostrados na figura 4.

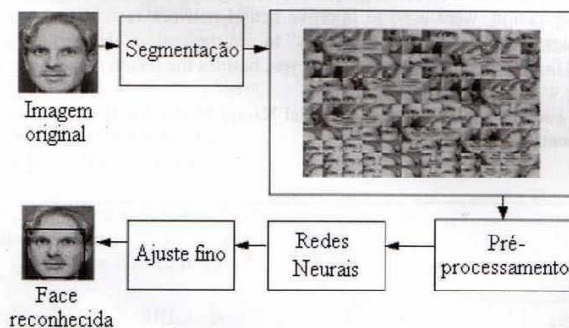


Figura 4 - Visão geral do sistema.

O primeiro módulo, denominado de segmentação, divide a imagem principal em várias sub-imagens quadradas de diferentes tamanhos, todas redimensionadas para um padrão de 19x19 pixels (*picture elements*), este módulo será detalhado na seção 3. O segundo módulo chamado de pré-processamento, descrito na seção 5, prepara adequadamente o sinal visando aumentar a capacidade de reconhecimentos corretos do terceiro módulo baseado em redes neurais artificiais treinadas com os bancos de dados semelhantes aos mostrados nas figuras 2 e 3.

As sub-imagens identificadas como faces pela rede neural são então separadas e apresentadas a um quarto módulo, o de ajuste fino, no qual um retângulo representativo da face é determinado e otimizado para se obter, finalmente, o menor retângulo que envolva a face de acordo com a definição mostrada na seção 1. Estes ajustes serão detalhados na seção 6.

3. SEGMENTAÇÃO

Partindo do princípio de que qualquer parte da imagem original, independentemente de tamanho ou proporção pode conter a face desejada, o módulo de segmentação deverá retirar, da imagem original, todos os retângulos possíveis e distintos enviando-os ao módulo seguinte redimensionados para um mesmo tamanho padrão, fixado em 19x19, para serem finalmente entregues aos módulos de pré-processamento e rede neural especializada no reconhecimento de padrões faciais.

Embora esse método seja possível de ser implementado, exigiria muito tempo de processamento, pois mesmo de uma imagem pequena podem-se extrair muitas diferentes sub-imagens retangulares distintas.

O método sugerido neste trabalho baseia-se em procurar as bordas da imagem original e segmentar a imagem apenas nas partes que contiverem pelo menos um ponto de borda.

A figura 5 mostra um exemplo desse processo no qual nota-se que a face a ser localizada certamente conterá mais do que um ponto de borda

Mesmo segmentando a imagem apenas nos pontos contendo bordas, o número de sub-imagens ainda pode ser bastante elevado.

Uma solução para minimizar esse problema é considerar inicialmente apenas sub-imagens quadradas, e posteriormente, fazer um ajuste fino (seção 6) nas imagens que forem reconhecidas como faces pela RNA.

Após a segmentação da imagem, todas as sub-imagens passam por um pré-processamento que será mostrado na seção 5.

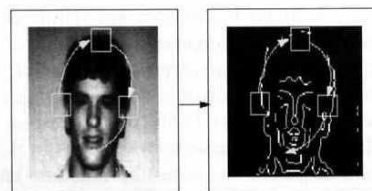


Figura 5 - Segmentação baseada nas bordas.

4. REDES NEURAIS ARTIFICIAIS

As principais características das redes neurais artificiais (RNAs) são a sua habilidade de aprender certo conhecimento e a sua capacidade de generalizar, adaptando-se a situações que jamais fizeram parte do conjunto de treinamento [5].

Uma RNA consiste em uma ou várias unidades processadoras (neurônios artificiais) básicas interligadas, dispostas em uma ou mais camadas.

Cada camada pode conter diversos elementos processadores. A conexão entre as unidades processadoras é feita através dos pesos sinápticos que, em outras palavras, definem a influência entre as unidades processadoras interligadas.

Estes pesos, conhecidos também como parâmetros internos da rede, têm seus valores atualizados durante o processo de treinamento. Nesse processo iterativo de absorção de conhecimento os pesos retêm a informação que se deseja que a rede aprenda.

Cada unidade processadora executa uma operação simples sobre a informação recebida da entrada externa ou da saída de outro processador [6].

O resultado é então transmitido para outros processadores que, por sua vez, repetem a operação. O processo de atualização dinâmica dos pesos é feito paralelamente e distribuído por toda a rede. Quando as saídas diferem dos alvos a serem atingidos, por um valor inferior à tolerância de erro preestabelecida, está concluído o processo de aprendizagem.

Neste ponto pode-se dizer que a rede está devidamente treinada e a matriz de pesos, obtida nesta etapa e que será usada na fase de execução, retém o conhecimento que a rede foi capaz de absorver durante a fase de aprendizagem.

A figura 6 mostra a arquitetura de uma rede neural genérica com uma camada, onde $p_1 \dots p_R$ representam os elementos do vetor de treinamento P da rede (estímulo externo). O bloco somatório realiza a soma ponderada dos pesos W_{ij} com as entradas $p_1 \dots p_R$ representada na forma matricial como $NET_j = W.P$

A saída $a_j = F(NE_j)$ é o resultado da aplicação da função de ativação F sobre a soma ponderada NET , sendo esta função linear ou não, com ou sem "bias" (b).

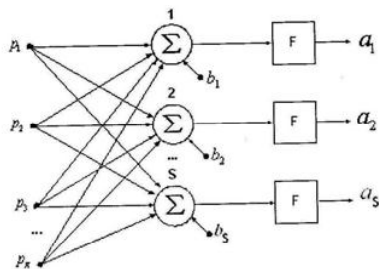


Figura 6 - Arquitetura genérica de uma RNA.

O processo de aprendizado envolve sucessivas apresentações dos padrões de treinamento à rede. Um processo iterativo de ajuste nos parâmetros internos (pesos), executado de modo a aproximar, dentro de limites preestabelecidos, a saída da rede a um alvo previamente fornecido.

A regra de *Widrow-Hoff*, mostrada no conjunto de equações 1, referente ao treinamento de Perceptrons lineares (ADALINE) [1], [4], é a base fundamental dos diversos métodos de treinamento

das RNAs, inclusive no algoritmo de retro-propagação do erro usado neste trabalho.

Ela avalia o erro médio quadrático a cada iteração através da derivada parcial do erro quadrático em relação ao peso W e ao limiar (bias) b , mostrados na figura 6.

Representando-se por t o vetor alvo a ser atingido no treinamento, P , o vetor de treinamento apresentado à rede, R , o número de elementos do vetor de treinamento, tem-se para o neurônio 1 na n -ésima iteração com $j = 1 \dots R$, o conjunto de equações (1).

$$\frac{\partial e^{2(n)}}{\partial W_{1j}} = 2e^{(n)} \frac{\partial e^{(n)}}{\partial W_{1j}}$$

$$\frac{\partial e^{2(n)}}{\partial b} = 2e^{(n)} \frac{\partial e^{(n)}}{\partial b}$$

$$\frac{\partial e^{(n)}}{\partial W_{1j}} = \frac{\partial [t^{(n)} - a^{(n)}]}{\partial W_{1j}} = \frac{\partial}{\partial W_{1j}} [t^{(n)} - (Wp^{(n)})]$$

onde:

$$Wp^{(n)} = \sum_{j=1}^R W_{1j} P_j^{(n)} + b$$

$$\frac{\partial e^{(n)}}{\partial W_{1j}} = -P_j^{(n)}$$

$$\frac{\partial e^{(n)}}{\partial b} = -1 \quad (1)$$

É mostrado, no conjunto de equações (2), a variação ΔW do peso e Δb do "bias", bem como a base do método usado no processo adaptativo de atualização dos parâmetros internos da rede [1], [4].

$$\Delta w^{(n)} = -2\eta e^{(n)} \frac{\partial e^{(n)}}{\partial W}$$

$$\Delta b^{(n)} = -2\eta e^{(n)} \frac{\partial e^{(n)}}{\partial b}$$

$$W^{(n+1)} = W^{(n)} + \Delta w^{(n)} = W^{(n)} + 2\eta e^{(n)} P^{(n)}$$

$$b^{(n+1)} = b^{(n)} + \Delta b^{(n)} = b^{(n)} + 2\eta e^{(n)} \quad (2)$$

No conjunto de equações (2), o termo η representa a taxa de aprendizado, que define a velocidade de convergência do método.

Assim, se η for muito grande o treinamento tende a ser feito mais rapidamente, porém existe a possibilidade de oscilação [1] e [4].

Por outro lado, se η for muito pequeno o treinamento pode ser lento e é possível que a rede fique paralisada ao atingir um mínimo local, sem que o treinamento evolua para o mínimo global.

Terminado o treinamento, a rede está pronta para ser executada. Nesta fase são fornecidos à RNA novos padrões, denominados padrões de teste, e a rede responde rapidamente, de acordo com o que lhe foi imposto durante o treinamento.

As RNAs possuem grande capacidade de tratar dados distorcidos por ruído. Um outro fato que merece destaque está relacionado à capacidade de poder considerar várias soluções possíveis para um problema simultaneamente, enquanto os tratamentos convencionais só podem avaliar uma solução de cada vez.

Segundo [1] e [4], estas características podem ser atribuídas à estrutura altamente paralela da rede, que permite a distribuição da funcionalidade da unidade processadora, ou seja, do neurônio artificial, por toda a rede, conforme ocorre na rede neural biológica.

5. PRÉ-PROCESSAMENTO

A rede neural proposta neste trabalho usa no treinamento cerca de 5000 exemplos de faces e não faces, semelhantes às mostradas nas figuras 2 e 3, e possui duas saídas, uma indicando a presença de um padrão facial, e outra indicando que não se trata de uma face. O algoritmo utilizado para o treinamento foi o *back-propagation* com gradiente conjugado [1].

Para se obter uma elevada taxa de reconhecimento correto é necessário realizar um processamento de sinais anterior à aplicação das imagens à RNA, chamado de pré-processamento [6].

No caso específico do reconhecimento de padrões faciais é necessário um processo que acentue as diferenças entre as imagens que contêm faces e as que não contêm [2].

Assim é necessário aplicar ao banco de dados de imagens disponível, uma sequência de transformações na imagem incluindo: equalização de histograma, filtros ovais, correção de iluminação e técnica de componentes principais [1].

A equalização de histograma é feita segundo os processos clássicos de equalização [3], com o objetivo de aumentar o contraste da imagem.

Já a utilização do filtro oval segue um princípio baseado na observação de que os pontos nas bordas da imagem não influenciam na classificação entre faces e não faces e, portanto, podem ser "mascarados".

Esta máscara evita que, durante o processo de treinamento e teste, estes pontos sejam levados em consideração pela RNA.

O filtro oval consiste em eliminar pixels pela aplicação de máscaras triangulares de tamanho ajustável aos cantos da imagem, conforme mostra a figura 7.



Figura 7 – "Mascaramento" por filtro oval.

A correção de iluminação é o ponto chave para um reconhecimento confiável e consiste em retirar o efeito de uma iluminação regular da imagem a ser reconhecida.

Supondo que a iluminação seja uma função linear, uma imagem na qual foi aplicado algum tipo de iluminação poderia ser decomposta em um gradiente linear somado a uma imagem isenta de iluminação.

Para encontrar esse gradiente linear a partir da imagem iluminada, propõe-se um método que calcula as médias dos valores de tons de cinza de todas as colunas da imagem e a seguir faz-se o mesmo com todas as linhas.

Para cada um desses dois vetores encontra-se, através do método dos mínimos quadrados mostrado nas equações (3), (4) e (5), a equação da reta $z = aw + b$ que melhor aproxima os valores dos vetores, onde z é o valor da média dos níveis de cinza, w a posição da

coluna ou linha na figura e a e b os coeficientes da reta.

Obtêm-se, através do cálculo de a e b nas equações (3), (4) e (5), as aproximações lineares dos vetores das médias das linhas e colunas da imagem.

$$z \cong aw + b \quad (3)$$

$$b = \frac{\sum z \cdot \sum w^2 - \sum w \cdot \sum w \cdot z}{z \cdot \sum w^2 - (\sum w)^2} \quad (4)$$

$$a = \frac{z \cdot \sum w \cdot z - \sum w \cdot \sum z}{z \cdot \sum w^2 - (\sum w)^2} \quad (5)$$

Com esses novos vetores, aproximados linearmente pelas médias das colunas e linhas, constrói-se uma máscara para reduzir o efeito da iluminação, assumida linear, a partir do produto externo dos mesmos.

Esta máscara é então subtraída da imagem original, corrigindo assim a presença de iluminação ambiente. A figura 8 ilustra esse processo.

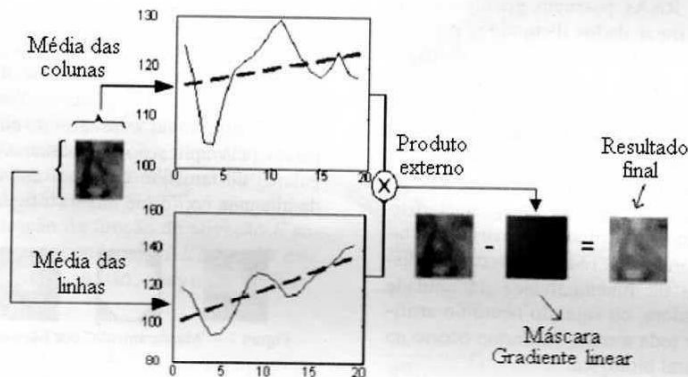


Figura 8 - Método de correção de iluminação.

A figura 9 mostra todas as transformações envolvidas na ordem em que as mesmas são aplicadas.



Figura 9 - Sequência de pré-processamentos.

A seguir forma-se uma matriz com todas as faces do banco de dados e aplica-se a técnica de componentes principais PCA [1], [8].

Esta técnica faz uma análise da variância do sinal a ser classificado, retirando os dados que tiverem baixa variância, permitindo uma redução em sua dimensionalidade.

ENGEVISTA, v. 5, n. 10, p. 19-30, dez. 2003

Assim, um vetor de dados X projetado em um vetor unitário q , de modo que $A = X^T q$, este poderá ser completamente reconstituído através de suas projeções, como mostra a equação (6), conhecida também como equação de síntese, na qual os a_j representam as componentes principais, e A o vetor projetado:

$$X = \sum_{j=1}^m a_j q_j \quad (6)$$

Porém, se ao invés de reconstituirmos o vetor X completamente com todas as suas projeções, truncarmos o somatório da equação (6) após n termos, um novo vetor X terá sua dimensão reduzida a n . Por causa desse truncamento.

A equação (7) mostra a relação entre o somatório de todas as variâncias das projeções eliminadas, na qual σ_j^2 representa a variância da j -ésima componente principal de a_j e λ_j os autovalores associados a cada uma das projeções. Assim, quanto mais próximo de zero forem os autovalores, mais eficiente será a redução de dimensionalidade [1].

$$\sum_{j=n+1}^m \sigma_{Eliminada}^2 = \sum_{j=n+1}^m \lambda_j \quad (7)$$

Portanto, se calcularmos anteriormente os m autovalores da matriz autocorrelação R de X e colocarmos os mesmos em ordem crescente, pode-se eliminar apenas as projeções dos menores autovalores, o que garantirá uma redução na dimensão de X com a menor perda possível de variância, e conseqüentemente de informação [1], [8].

Vale ressaltar que, neste trabalho, a geração da matriz de transformação para os componentes principais é feita apenas com base no banco de dados de faces (figura 2).

No método de segmentação, mostrado na seção 3, uma face tende a ser identificada várias vezes em quadrados adjacentes, como mostra a figura 10.

Isso não chega a ser uma desvantagem, pois estas múltiplas detecções podem ser usadas para validar e encontrar a correta localização da face. A redução por centróide reduz esse problema.

Inicialmente, define-se como “vizinhos de centróide” quando um quadrado possui seu centro de massa dentro da área limitada por outro quadrado e este outro tem seu centro de massa dentro da área limitada pelo primeiro. Esta definição pode ser estendida para n quadrados, o que é mostrado na figura 10, onde todos os quadrados mais claros são “vizinhos de centróide”.

Neste trabalho, durante a redução por centróide, os “vizinhos de centróide” serão reduzidos a apenas um quadrado cujo tamanho será a média dos respectivos vizinhos. Em contrapartida, os quadrados que não possuírem pelo menos um “vizinho de centróide” podem ser desconsiderados, pois os mesmos serão, muito provavelmente, erros do sistema de detecção ou da RNA, já que faces são sempre detectadas várias vezes.

A figura 10 mostra uma redução por centróide, na qual fica evidente que o quadrado superior direito, mais escuro, não entra no cálculo da média, pois o mesmo não possui nenhum “vizinho de centróide”.

6. AJUSTE FINO

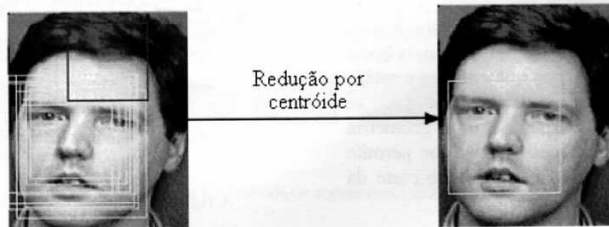


Figura 10 - Redução por Centróide.

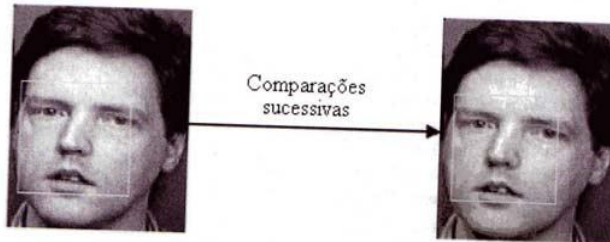


Figura 11 - Redução por Comparações Sucessivas.

Além da redução por centróide, é feito também um outro ajuste fino chamado "redução por comparações sucessivas".

A partir do quadrado de saída da redução por centróide são gerados quadrados deslocados de um ponto à direita, à esquerda, acima e abaixo, além de quadrados com tamanho reduzido em também um ponto. Esses quadrados são checados novamente com a rede neural, fazendo-se uma "busca gulosa" em um processo recursivo, na qual se assume como face atual o maior valor de saída da rede neural. Esse processo se repete até que não se tenha encontrado um valor maior que o atual.

A figura 11 mostra o ajuste por comparações sucessivas. Nota-se uma melhora na detecção, que transformou o quadrado em um retângulo com largura um pouco menor, mais adaptado à face.

7. LOCALIZAÇÃO DOS ELEMENTOS DA FACE

Uma vez localizada a face, com os métodos descritos nas seções 3, 5 e 6, pode-se desenvolver técnicas simples na localização dos olhos, do nariz e da boca. Isso é possível, pois, obtendo-se o retângulo mínimo que envolve a face, pode-se dizer que há certamente uma geometria própria dos padrões faciais, que permite segmentar e localizar cada sub-parte da face, como olhos, nariz e boca.

7.1 LOCALIZAÇÃO DOS OLHOS

Obtendo-se o retângulo mínimo que envolve a face é possível através da geometria facial, garantir com grande probabilidade, que os olhos estarão na parte superior deste retângulo mínimo. É possível ainda se admitir que caso a face tenha sua metade superior dividida em duas partes iguais, garante-se, com elevado nível de certeza, que os olhos, direito e esquerdo, estarão em suas respectivas metades.

Para se obter a correta localização dos olhos, calcula-se a correlação cruzada bidimensional $f(x, y) \circ g(x, y)$ [3] mostrada na equação (9) entre a metade superior direita ou esquerda escolhendo-se o padrão da figura 12 como o mais semelhante aos vários olhos disponíveis no banco de dados. Na equação (9), M e N são os tamanhos horizontal e vertical da imagem.

$$f(x, y) \circ g(x, y) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(x, y)g(x+m, y+n) \quad (9)$$



Figura 12 - Padrão dos olhos.

A figura 13 ilustra esse processo:

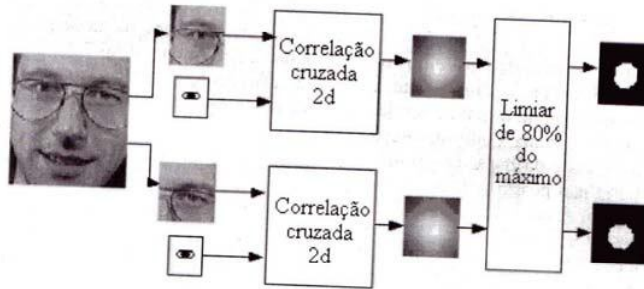


Figura 13 - Método de correlação cruzada.



Figura 14 - Posição final dos olhos na face.

Obtendo-se as duas máscaras, pode-se colocá-las lado a lado para encontrar a posição final dos dois olhos, como mostra a figura 14.

sando o método Laplaciano [3], conforme a equação (10).

$$\nabla^2 f(x, y) = \frac{\partial^2}{\partial x^2} f(x, y) + \frac{\partial^2}{\partial y^2} f(x, y) \quad (10)$$

7.2 LOCALIZAÇÃO DA BOCA

A posição da boca pode ser encontrada ainda mais facilmente que os olhos.

Inicialmente divide-se a imagem em duas partes iguais na horizontal. A seguir, encontram-se as bordas desta imagem u-

A seguir, ignorando-se as bordas laterais, faz-se uma segmentação procurando-se a maior região conexa disponível e que será muito provavelmente a boca nas imagens de faces investigadas. A figura 15 ilustra esse processo:

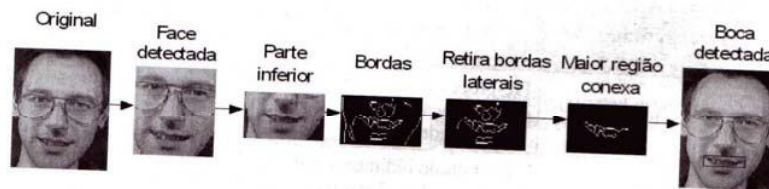


Figura 15 - Método de segmentação da boca.

7.3 LOCALIZAÇÃO DO NARIZ

Embora a posição da boca e dos olhos possa ser detectada separadamente, a posição do nariz é muito difícil de ser detectada sem que se tenha a informação da posição dos olhos e da boca. Isto ocorre porque o nariz não possui muitas bordas detectáveis, nem é, na maioria dos casos, mais escuro ou mais claro que as outras partes da face tornando, conseqüentemente, sua detecção menos confiável.

Entretanto, possuindo-se previamente a posição dos olhos e da boca, descritas nas seções 7.1 e 7.2, pode-se garantir, para faces rotacionadas até ± 45 , que a posição do nariz estará no quadrado definido pela posição dos centróides dos olhos e entre o centróide da boca e a menor coordenada *Y* dos centróides dos olhos, conforme mostra a figura 16.

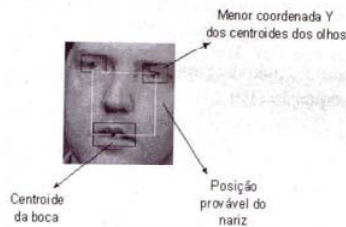


Figura 16 - Posição do nariz.

De posse desta região provável, segmenta-se o nariz da mesma forma que a boca, descrita na seção 7.2, eliminando-se novamente as bordas laterais e procurando-se a maior região conexa.

8. RESULTADOS

Implementando-se as técnicas de localização facial, descritas nas seções 3, 5 e 6, em conjunto com os métodos de localização de olhos, boca e nariz (seções 7.1, 7.2 e 7.3) constrói-se o localizador de padrões faciais.

Usando o localizador proposto, foram executados 400 testes de imagens de faces disponibilizadas na página do laboratório da AT&T [7]. Esse banco de dados de imagens, embora não possua quase nenhuma variação na distância focal entre a câmera e o indivíduo, possui faces ligeiramente rotacionadas e poses não totalmente frontais. Estas características são fundamentais para testar o sistema desenvolvido, visto que ainda é um desafio localizar faces corretamente nesse tipo de padrão facial.

Os resultados preliminares mostrados na tabela 1 indicam que a maior dificuldade é segmentar e extrair a posição do nariz, pois o mesmo normalmente não possui bordas que o diferenciem de outras partes do rosto e pode ser afetado pela presença de padrões com óculos e barba, por exemplo.

Tabela 1 - Resultados encontrados e técnicas empregadas

Padrão	Técnicas empregadas	Taxa de Acertos
Face	Pré-processamentos, correção de iluminação, segmentação e Redes neurais	90%
Olhos	Correlação bidimensional com padrão dos olhos	83%
Nariz	Deteção de bordas e Crescimento de regiões conexas	68%
Boca	Deteção de bordas e Crescimento de regiões conexas	86%

A figura 17 mostra várias saídas do sistema, em linha sólida tem-se a saída relativa à detecção neural da face. Em pontilhado tem-se a segmentação da bo-

ca, em tracejado escuro tem-se a segmentação dos olhos e finalmente, em tracejado claro, a segmentação do nariz.

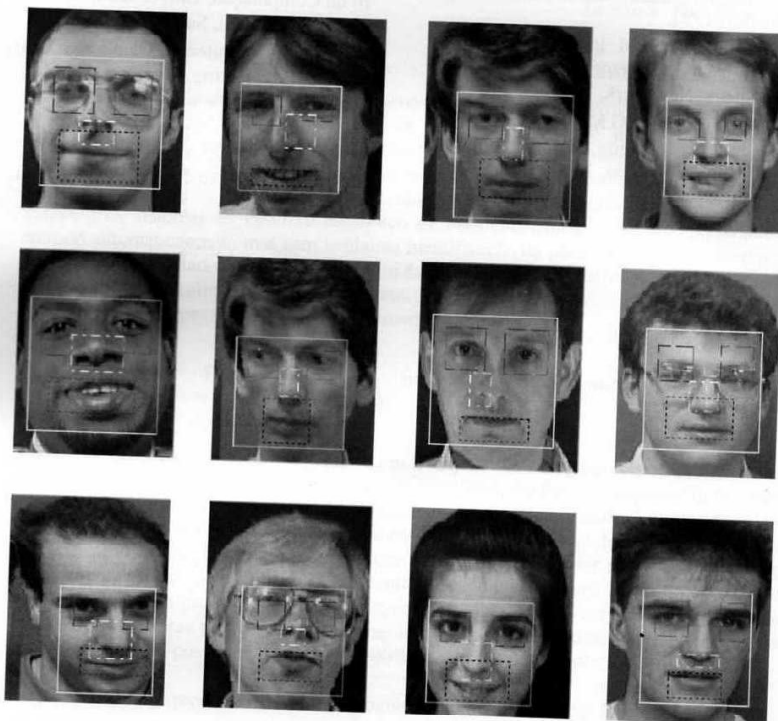


Figura 17 - Resultados do Sistema Localizador de Padrões Faciais.

9. CONCLUSÕES E TRABALHOS FUTUROS

Pode-se ver, pelos resultados da tabela 1 e pela figura 17, que o Sistema Localizador de Padrão Facial tem possibilitado resultados animadores, conseguindo localizar na imagem original a face, os olhos e a boca do indivíduo, sendo ainda difícil localizar a posição do nariz. Esse problema deve-se ao fato de que o nariz não possui bordas nem nenhuma característica que o diferencie do resto da face.

Em função dos resultados obtidos, pode-se dizer que a taxa de acertos poderia ser em muito aumentada, caso se pudesse garantir que as imagens dos indivi-

duos estivessem sempre de frente e sem nenhum tipo de rotação.

No momento estão sendo realizadas novas pesquisas no sistema de segmentação e extração de características visando melhorar a eficiência total do localizador, sendo também possível usar os resultados deste trabalho em um sistema de identificação de indivíduos.

10. REFERÊNCIAS

- [1] HAYKIN S.; Neural Networks, a comprehensive Foundation; 2a edition. Prentice Hall Press, New Jersey, 1999.
- [2] ROWLEY H. A.; Neural Network-Based Face Detection; School of Computer Science, Computer Science De-

partment, Carnegie Mellon University Pittsburgh, 1999.

[3] GONZALEZ R. C. and WOODS R. E.; Digital Image Processing, Addison-Wesley, 1992.

[4] ZURADA, J. M. Introduction to Artificial Neural Systems, West Publishing Company, New York, 1992.

[5] SHEPHERD, G.M.; KOCH. C. The synaptic organization of the brain. Oxford University Press, New York, 1990.

[6] MOUTINHO, A. M.; BIONDI, L. N. Métodos de Pré-Processamento de Sinais Aplicados Ao Treinamento De Redes Neurais Artificiais. II Congresso Brasileiro de Computação, Universidade do Vale do Itajaí - Univali, Santa Catarina, 2002.

[7] AT&T Laboratories, Cambridge University Engineering Department, Cambridge, <http://www.uk.research.att.com>, 2003.

[8] JOHNSON, R. A. Applied Multivariate Statistical Analysis, Prentice-Hall, EUA, 1998.