ALGORITMO PARA DETECÇÃO DE BOCA EM FACES HUMANAS USANDO MATRIZ DE CO-OCORRÊNCIA E SVM*

ALGORITHM FOR DETECTION OF MOUTH IN HUMAN FACES USING MATRIX OF CO-OCCURRENCE AND SVM

ALGORITMO PARA LA DETECCIÓN DE BOCA EN ROSTROS HUMANOS USANDO LA MATRIZ DE CO-EXISTENCIA Y SVM

Tarcísio Sousa de Almeida
Eduardo da Cunha Batista
Eduardo de Jesus Coelho Reis
Artur Bernardo Silva Reis
Aristófanes Corrêa Silva

Resumo: Esse artigo propõe uma metodologia computacional capaz de detectar a componente facial boca em imagens de faces frontais humanas submetidas a diversas condições de iluminação, textura, tamanho, etc. Para a obtenção dos resultados, foram utilizadas técnicas de Processamento de Imagens e Visão Computacional. Para classificá-las, são usados atributos, advindos dos descritores de Haralick, que descrevem as texturas das imagens. A SVM (Máquina de Vetor Suporte, do inglês Support Vector Machine) é usada para reconhecer os padrões das texturas das imagens utilizadas. Os resultados preliminares se mostraram bem promissores, pois atingiram cerca de 84% de acerto na classificação da boca, levando-se em conta que essa metodologia é generalista.
Palavras-chave: Detecção de bocas. Matriz de co-ocorrência. Descritores de Haralick. Maquina de vetores suporte.

Abstract: This paper proposes a computational methodology that is able to detect the facial component mouth in frontal human face images submitted to a lot of conditions like light, texture, size, etc. To get the results, techniques of Image Processing and Computational Vision were used. Attributes that describe were used to classify the textures of the images, these attributes are obtained by the Haralick descriptors. A SVM (Support Vector Machine) is used to recognize patterns of the textures from the images. The preliminary results were very promising, because the hit rate of mouth classification reached 84% that is a good result considering that this is a generalistic methodology.
Keywords: Detection of mouths. Co-occurrence matrix. Haralick descriptors. Support vector machine.

Resumen: Este trabajo propone un método computacional capaz de detectar la boca en imágenes faciales de la corteza frontal humana expuestas a diversas condiciones de iluminación, textura, tamaño, etc. Para la obtención de los resultados fueron utilizadas técnicas de Procesamiento de Imágenes y de Visualización Computacional Para clasificarlas, fueron usados los atributos provenientes de los descriptores de Haralick. La SVM (Máquina de Soporte Vectorial, del Inglés Support Vector Machine,) se utiliza para reconocer los padrones de las texturas de las imágenes utilizadas. Los resultados preliminares han demostrado ser prometedores ya que llegaron a aproximadamente el 84% de clasificación correcta de la boca que es un buen resultado teniendo en cuenta que esta metodología es de carácter general.
Palabras clave: Detección de la boca. Matriz de co-existencia. Descriptores de Haralick. Máquinas de vectores soporte.

1 INTRODUÇÃO

Reconhecimento Facial consiste na atividade de identificar as faces presentes em imagens ou vídeo de uma cena. Esta área de pesquisa científica tem se destacado devido a suas inúmeras aplicabilidades comerciais em áreas como segurança, biometria ou em centros de arquivo de fotos de personalidades.

O ato de reconhecer uma face envolve uma série de dificuldades, como a posição da face, expressão facial, oclusão de elementos faciais

*Trabalho premiado durante o XXIII Encontro do SEMIC realizado na UFMA entre os dias 08 a 11 de novembro de 2011. O artigo foi produzido a partir de pesquisas realizadas no LabPAI (Laboratório de Processamento de Imagem) da UFMA.
*Artigo recebido em maio 2012
Aprovado em agosto 2012

importantes, como algum dos olhos ou a boca, dificultando a atividade proposta e acarretando a necessidade de algoritmos robustos.

Uma importante aplicação para reconhecer face encontra-se na organização de grandes arquivos de dados, onde são necessárias pessoas altamente especializadas. A pessoa, não estando mais disponível, perde-se a expertise de busca às informações arquivadas. No arquivamento de imagens de personalidades, este fato se torna especialmente crítico, uma vez que uma determinada pessoa se "familiariza" com personagens de um dado período e passa a reconhecê-los em fotografias ainda não catalogadas, tarefa que seria impossível de ser executada por pessoa não familiarizada com tais personagens.

Tendo em vista a problemática observada, nota-se a necessidade de um suporte para o reconhecimento de pessoas ao profissional que manipula as imagens do arquivo de personalidades.

No entanto, para que ocorra o reconhecimento facial, é imprescindível que antes sejam encontradas as faces presentes na imagem. Seguindo os métodos baseados no conhecimento, a procura facial ocorreria através da detecção de elementos inerentes à face. Munido destas informações, este trabalho descreve uma metodologia ágil para detectar a componente facial boca. A detecção de boca em faces humanas será uma ferramenta útil no processo de reconhecimento de personalidades históricas, uma vez que, além de restringir a área de busca apenas à face, ele terá alta abrangência por ser a boca um elemento inerente à face.

Este artigo tem por objetivo apresentar os resultados preliminares de uma metodologia computacional capaz de detectar, com um bom nível de precisão, a componente facial boca em imagens de faces humanas, submetidas a diversas condições de iluminação, textura e escala, usando Técnicas de Processamento de Imagens e Visão Computacional.

Este trabalho faz parte de estudos que estão sendo realizados no Laboratório de Processamento e Análise de Imagens (LabPAI) da Universidade Federal do Maranhão que desenvolve uma ampla pesquisa em rede denominada: "Reconhecimento de face de personalidades históricas para auxiliar na recuperação de documentos em acervos digitais", com a colaboração do Centro de Matemática Aplicada (CMA-FGV) e do Centro de Pesquisa e Documentação de História Contemporânea (CPDOC-FGV), ambas integrantes da Fundação Getúlio Vargas (FGV).

Este artigo está subdividido em quatro seções. Na Seção 2, será dado enfoque sucinto à fundamentação teórica, ou seja, aos métodos de extração de características (matriz de co-ocorrência) e ao método de sua classificação (Máquina de Vetor de Suporte - SVM). Na Seção 3, será apresentada a metodologia proposta. Os resultados e discussões serão apresentados na Seção 4. Por fim, na Seção 5, serão feitas as conclusões finais.

2 FUNDAMENTAÇÃO TEÓRICA

Após a análise de algumas técnicas de detecção de imagens, optou-se por utilizar a matriz de co-ocorrência por se mostrar uma técnica eficiente, apesar de computacionalmente cara, uma vez que para cada imagem cria-se uma matriz quadrada de ordem igual ao maior nível de cinza desta. A matriz gerada contêm todas as passagens de níveis de cinza da imagem que posteriormente serão usadas para aquisição das probabilidades de cada ocorrência.

2.1 Matriz de Co-ocorrência

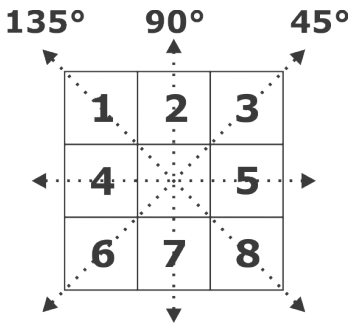
A matriz de co-ocorrência, proposta por Haralick (HARALICK; SHANMUGAM, 1973), consiste num método para estimar as combinações de diferentes intensidades de cinza na imagem, possibilitando a caracterização das texturas nela presentes através de um conjunto de estatísticas de ocorrências de níveis de cinza. Seus parâmetros texturais são considerados como parâmetros estatísticos de segunda ordem por não serem derivados diretamente de dados das imagens (PHAM; ALCOCK, 1998). A ideia inicial do reconhecimento de texturas, assim como outras atividades da área de análise de imagens, consiste em extrair da imagem algumas características que permitam realizar posteriormente um discernimento, uma tomada de decisão ou uma classificação.

Cada elemento da matriz conterá o número de ocorrências de combinações entre um pixel analisado e o pixel vizinho, dado uma distância d e a uma orientação θ, quer seja vertical, diagonal ou horizontal.

A matriz é bidimensional e obrigatoriamente quadrada, com o número de linhas e colunas determinado pela quantidade de níveis de cinza presentes na imagem. Portanto, quanto mais níveis de cinza, maior a

matriz. A matriz de co-ocorrência, na sua forma geral, pode ser especificada por uma matriz de frequências relativas P (i, j, d, θ), ou seja, a probabilidade de ocorrer um par de pixels i, j, separados por uma distância d em uma orientação θ, numa determinada imagem, pode ser visto na Figura 1.

Figura 1 - Orientações possíveis da matriz de co-ocorrência



Fonte: (HARALICK; SHANMUGAM, 1973)

Depois de explicados o conceito e os parâmetros da matriz de co-ocorrência, sua construção pode ser iniciada. Uma imagem com N níveis de cinza é escolhida. A Figura 2 utilizada como exemplo apresenta três níveis de cinza: $N_0 = 0; N_1 = 1; N_2 = 2$.

Figura 2 - Imagem com 3 níveis de cinza

0	0	0	1	2
1	1	0	1	1
2	2	1	0	0
1	1	0	2	0
0	0	1	0	1

Fonte: (RAFAEL et al., 2002.)

Para extrair as características da imagem, são geralmente escolhidas quatro direções: 0°, 90°, 180°, 360°, e a distância adotada entre os pixels é de 1. Por exemplo, consideremos que a distância definida seja 1 e a orientação 0º. Na posição (0,0) da matriz (Figura 2), haverá o número de vezes em que ocorreu a passagem do nível de cinza 0 para outro pixel imediatamente a sua direita contendo mesmo valor. Na imagem podemos ver que esta combinação é encontrada 4 vezes. Na posição (0,1) da matriz (Figura 2), será armazenado o número de vezes em que o nível 0 e o nível 1 encontram-se lado a lado, isto é, nível de cinza 0, distância 1 e orientação 0º. Assim prosseguiremos em cada posição da matriz até que ela seja totalmente preenchida.

Depois de construída, a matriz de co-ocorrência é normalizada, para enfim se descobrir a probabilidade de determinada transição de um nível de cinza para outro. Conforme a Equação 1 (onde M é a matriz de co-ocorrência de ordem N e i, j os índices), a matriz normalizada $P_{i,j}$ é gerada através da divisão de cada posição da matriz pelo somatório de todos os valores contidos nela.

$$P_{i,j} = \frac{M_{i,j}}{\sum_{i,j=0}^{N-1} M_{(i,j)}} \qquad (1)$$

Cada posição da matriz normalizada (Figura 3) é representada por $P_{i,j}$. O somatório de todos os valores deve ser um, correspondendo a 100%.

Figura 3-Exemplo de matriz de co-ocorrência normalizada

	0	1	2
0	0,25	0,125	0,0
1	0,125	0,1875	0,125
2	0,0625	0,125	0,0

Fonte: Elaborada pelos autores, com base na figura 2

Com a matriz de co-ocorrência, é possível adquirir descritores estatísticos da imagem denominados descritores de Haralick (HARALICK; SHANMUGAM, 1973), como entropia, correlação, contraste e variância. Alguns são exemplificados a seguir.

2.1.1 Correlação

Retorna uma medida de quão correlacionado está um pixel com o seu vizinho. A comparação é realizada em todos os pixels da imagem e a faixa de valores possíveis está entre -1 a 1. A correlação é 1 para duas imagens totalmente idênticas ou -1 para duas completamente descorrelacionadas. É formulada de acordo com a Equação 2, sendo i o número da linha, j o da coluna e $P_{i,j}$ a posição da matriz normalizada, μ_x e μ_x são as médias dos valores dos pixels nas direções x e y e σ_x e σ_x os seus desvios-padrão.

$$\sum_{i,j} \frac{ij P_{i,j} - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (2)$$

2.1.2 Contraste

Retorna medida de contraste entre a intensidade de um pixel analisado e seu pixel vizinho. Medidas relacionadas a contraste usam pesos relacionados com a distância da diagonal da matriz de co-ocorrência. Valores na diagonal da matriz representam pouco contraste. O contraste é caracterizado de acordo com a Equação 3, sendo i o número da linha, j o da coluna e $P_{i,j}$ a posição da matriz normalizada.

$$\sum_{i,j} \frac{|i-j|^2 P_{i,j}}{\sigma_x \sigma_y} \quad (3)$$

2.1.3 Homogeneidade

A homogeneidade pode ser descrita pela Equação 4, sendo i o número da linha, j o da coluna e $P_{i,j}$ a posição da matriz normalizada, como um valor que representa a proximidade da distribuição dos elementos em relação à diagonal da matriz de co-ocorrência dos níveis de cinza. Varia de 0, ausência de homogeneidade, a 1, que representa uma imagem completamente homogênea, uma matriz diagonal de co-ocorrência de níveis de cinza.

$$\sum_{i,j} \frac{P_{i,j}}{1+|i-j|} \quad (4)$$

2.1.4 Segundo Momento Angular

Descrito pela Equação 5, onde $P_{i,j}$ é a posição da matriz normalizada, mede o nível de homogeneidade da imagem, ou seja, a quantidade de repetições de pares de pixels na imagem.

$$\sum_{i,j} (P_{i,j})^2 \quad (5)$$

2.1.5 Entropia

A Entropia, descrita pela Equação 6, mede a desordem da imagem. Portanto, essa medida assume valores elevados quando a imagem possui textura não uniforme. A entropia está fortemente relacionada, porém de maneira inversa, com o segundo momento angular (BARALDI; PARMIGGIANI, 1995).

$$-\sum_{i,j} P_{i,j} \log(P_{i,j}) \quad (6)$$

2.1.6 Variância

Descrito pela Equação 7, onde i é o número da linha, j, o da coluna, $P_{i,j}$, a posição da matriz normalizada e μ, a média, é uma medida da heterogeneidade da textura. O valor da variância aumenta quando os valores dos níveis de cinza diferem de sua média.

$$\sum_{i,j} (i - \mu)^2 P_{i,j} \quad (7)$$

2.2 Máquina de Vetores de Suporte (SVM)

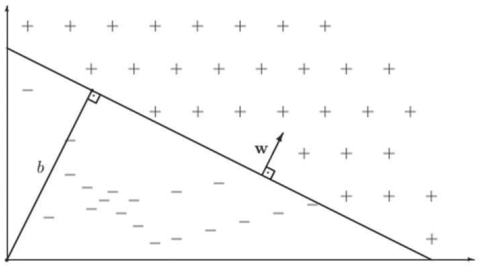
A Máquina de Vetores de Suporte (CORTES; VAPNIK, 1995) ou SVM , do inglês: *Support Vector Machine*, tem como intuito resolver problemas de classificação de padrões. É um método de aprendizado supervisionado que analisa os dados e reconhece padrões de acordo com as bases de treinamento que possui. As SVM buscam criar um hiperplano como superfície de decisão, cuja separação entre os exemplos dados seja máxima. A margem é obtida pela distância entre o hiperplano e os vetores mais próximos a ele, denominados vetores de suporte. Inicialmente, as SVM eram restritamente implementadas para conjuntos de dados de treinamento linearmente separáveis. Mais tarde, os resultados foram estendidos para dados não separáveis perfeitamente. Duas classes são ditas linearmente separáveis caso seja possível separar os dados em duas classes por um hiperplano, que é um plano de dimensão qualquer, ou seja, é a generalização, para o espaço n-dimensional, do conceito da reta.

Seja $f(x) = (w \cdot y) + b$ um hiperplano, onde w e b, obtidos no processo de aprendizagem a partir dos dados de entrada, conhecidos como *vetor de peso e bias*, responsáveis por controlar a função e a regra de decisão, podemos definir margem como a menor dis-

tância entre os exemplos do conjunto de treinamento e o hiperplano utilizado para separação destas classes.

Como exemplo podemos citar uma reta (Figura 4) como hiperplano, onde a linha diagonal corresponde ao hiperplano com a classe positiva acima representada por uma cruz e a classe negativa abaixo denotada por um traço. O vetor $w \in \mathbb{R}$ define uma direção perpendicular ao hiperplano. Enquanto variar o valor de $b \in \mathbb{R}$,o hiperplano move-se paralelamente a ele mesmo.

Figura 4-Espaço vetorial contendo as duas classes linearmente separáveis



Porém, em muitos problemas, as classes de dados não são linearmente separáveis. Para tais casos, usa-se as funções de *kernel* (VAPNIK, 1998.), responsáveis pela mudança de dimensionalidade das classes de dados, transformando-as em classes linearmente separáveis.

Para um bom entendimento sobre as funções de Kernel é necessário um prévio conhecimento sobre espaço de características. Em resumo um espaço de características é um espaço de alta dimensionalidade onde serão mapeados os conjuntos de dados de entrada que por intermédio de uma função de Kernel transforma este espaço em outro espaço de alta dimensionalidade, onde o conjunto de dados é linearmente separável.

As funções de Kernel, são definidas a partir do produto interno entre dois valores de entrada no espaço de características, de acordo com a Equação 5:

$$\mathrm{K}(\boldsymbol{x}_i,\boldsymbol{x}_j) = \langle \Phi(\boldsymbol{x}_i) \;\; \Phi(\boldsymbol{x}_j) \rangle \quad \textbf{(8)}$$

Existem vários de tipos de Kernel, definidos a partir de funções de *kernel* específicas, mas neste trabalho será utilizada a função Gaussiana, descrita pela Equação 6, onde x_i, x_j são valores de entrada no espaço de características e σ é o desvio-padrão.

$$\exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad \textbf{(9)}$$

Maiores informações podem ser encontradas em Vapnik (1998).

3 METODOLOGIA PROPOSTA

A metodologia tem por objetivo a detecção da componente facial boca, de forma que auxilie na diminuição de erros na validação da face e no seu posterior reconhecimento. Escolheu-se a opção de restringir o campo de busca da boca às faces, de forma a aumentar o número de acertos, elevando o nível de confiabilidade da metodologia.

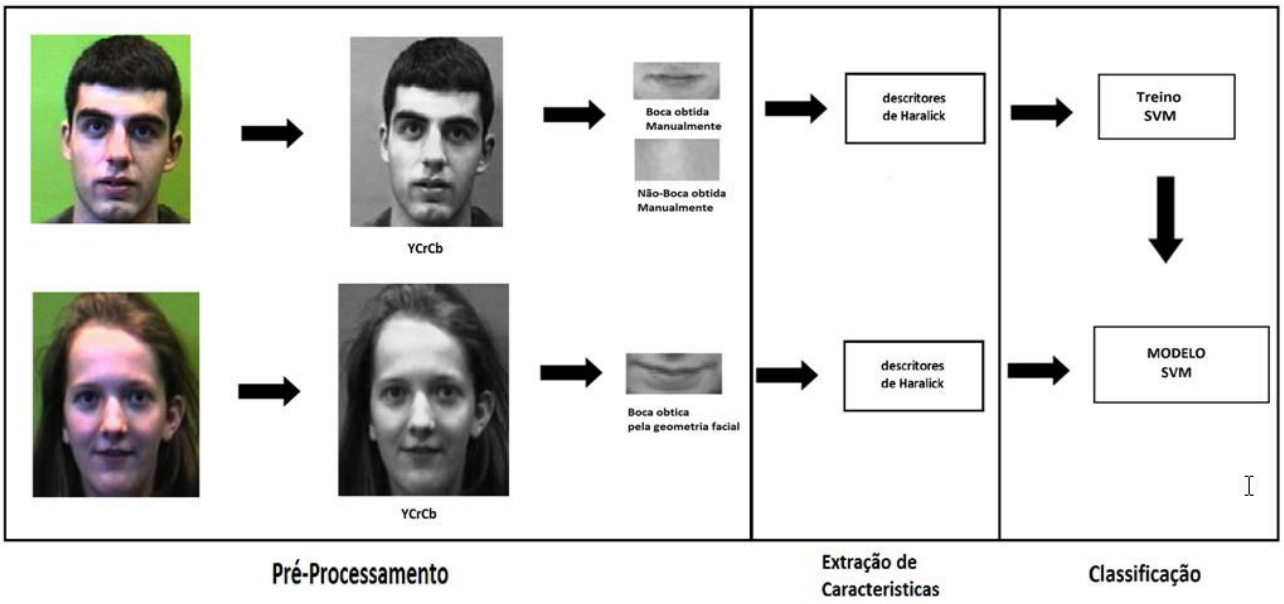
A metodologia adotada foi dividida nas seguintes etapas: Pré-processamento,Extração de Características e Classificação. A Figura 5 resume todos os passos da metodologia proposta.

3.1 Pré-Processamento

Neste passo, a imagem selecionada é convertida para nível de cinza, utilizando o modelo YCbCr, já que a técnica usada para extração de características é a matriz de co-ocorrência . Neste modelo, a informação de luminância é representada pelo componente Y, enquanto a informação de cor é armazenada nos componentes Cb e Cr (PEDRINI; SCHWARTZ, 2008). A imagem é convertida do sistema RGB para o componente Y do modelo YCbCr, conforme a Equação 10, pois não utilizará as cores armazenadas em Cb e Cr.

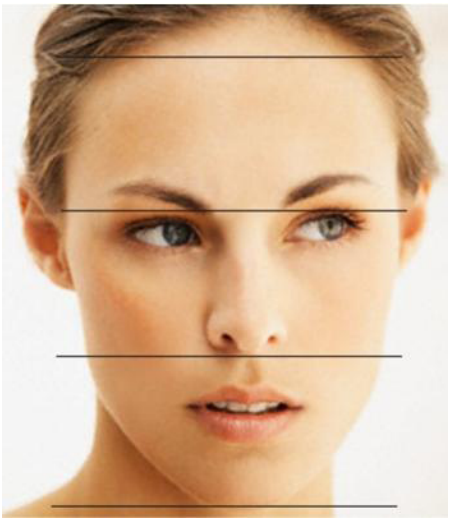
$$Y = 0.299 * R + 0.587 * G + 0.114B \quad \textbf{(10)}$$

É importante ressaltar que este passo foi dividido em duas partes: a manual, onde foi selecionada uma base de imagens de boca para criar um modelo que reconheça um padrão nessas imagens; e a automática, onde o software VisualFace, criado pela pesquisa (Reconhecimento de face de personalidades históricas para auxiliar na recuperação de documentos em acervos digitais), detectava supostas faces humanas, convertidas para escala de cinza. Por fim, é feita a detecção da possível região da boca através da geometria facial. A justificativa para a aplicação desse método é que faces frontais tendem a ter um formato padrão, descrito pela Figura 6.



Fonte: Elaborada pelos autores

Figura 6 - Padrão geométrico facial



Fonte: Google Imagens

3.2 Extração de Características

A extração das características das texturas presentes na imagem foi realizada através do uso da matriz de co-ocorrência. A matriz é gerada, contendo todas as passagens de níveis de cinza da imagem para, posteriormente, adquirir as probabilidades de cada ocorrência.

Primeiro, foram levantadas de cada uma das imagens de boca seis características do total de dezesseis, denominadas descritores de Haralick, a saber: contraste, entropia, segundo momento angular, correlação, variância e homogeneidade, obtidos através da matriz de co-ocorrência. Tais descritores são responsáveis por caracterizar a textura das imagens. Essas informações são armazenadas em vetores. O mesmo processo é repetido com as imagens que não eram boca. Na Figura 7, há exemplos de bocas e não bocas utilizados neste passo.

Figura 7 - Bocas e Não bocas

BOCAS	NÃO-BOCAS
	

Fonte: Elaborada pelos autores

A matriz de co-ocorrência da suposta região da boca, selecionada através da geometria facial, é gerada conforme descrito na fundamentação teórica. Os descritores fornecidos pela matriz são submetidos à SVM, responsável pela classificação.

3.3 Classificação

Para fazer o aprendizado da SVM, foram criadas duas classes: boca e não boca. A primeira continha imagens de bocas, enquanto a segunda de elementos que possivelmente não eram boca, como, por exemplo, texturas de pele humana.

Com a criação dessas duas classes, foi realizado o aprendizado da SVM, de acordo com os descritores obtidos pela matriz de co-ocorrência da imagem de entrada, à qual seria classificada de acordo com as duas classes previamente criadas, gerando, assim, um arquivo que contivesse todos os descritores de todas as imagens (bocas e não bocas).

4 RESULTADOS

Foram utilizadas 10432 imagens de bocas e de não bocas para o treinamento do SVM, sendo 1000 de bocas e 9432 não bocas. O motivo da existência de um número maior de imagens de não bocas reside no fato de que, na face humana, só existe uma boca, logo todas as outras partes são de não bocas.

A classificação feita pela SVM em todas 10432 imagens obteve cerca de 84% de precisão, ou seja, cerca de 84% das imagens foram classificadas da forma correta (tanto as bocas quanto as não bocas). A partir dos resultados apurados, observa-se que a detecção de boca em faces humanas será uma ferramenta útil no processo de reconhecimento de personalidades históricas, uma vez que, além de restringir a área de busca apenas à face, ele terá alta abrangência por ser a boca um elemento inerente à face.

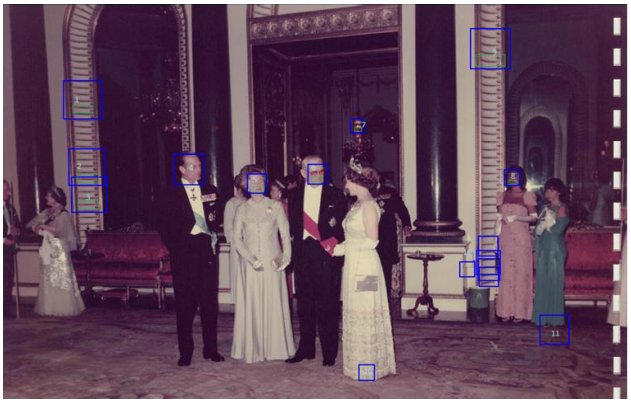
Figura 8 - Detecção de bocas em faces frontais



Fonte: Elaborada pelos autores

A Figura 8 mostra um exemplo onde detecção de bocas é falha. O contorno azul faz referencia a possíveis faces que foram marcadas pelo VisualFace produzido com o trabalho. O contorno verde dentro do contorno azul faz referência as supostas bocas encontradas com a nossa metodologia. Alguns contornos azuis foram numerados manualmente a fim de ilustrar os casos positivos e negativos. Os casos 1, 2, 3, 7 e 9 exibidos na imagem falharam devido à semelhança com boca. Os casos 10,11 e aglomerado de contornos azuis 12 da imagem, não houve detecção de boca, uma vez que o programa (VisualFace) detecta uma suposta "face", a não detecção da boca, descarta aquela "face" como falsa, logo, o programa mostrou-se eficiente. O caso 4 mostra um típico caso onde a face não está frontal, logo não houve detecção. Já os casos 2 e 3 houve sucesso.

Figura 9 - Exemplo onde a detecção da boca falha



Fonte: Elaborada pelos autores

5 CONSIDERAÇÕES FINAIS

Os resultados obtidos através dessa metodologia se mostraram bem promissores, tendo em vista que o trabalho está em fase inicial e ainda pode ser aperfeiçoado. Uma taxa de acerto média de 84% pode parecer ineficiente se comparado a outras metodologias propostas. A vantagem desse método em relação aos outros é que foi realizado para imagens sujeitas a qualquer condição e tamanho, diferente de outros métodos que utilizam uma câmera específica para uma aplicação específica. Esse método é generalista e tem como principal objetivo fazer a detecção da boca em imagens antigas escaneadas.

Aperfeiçoamentos desse método devem ser feitos para obtenção de melhores resultados e maior agilidade na sua execução. Atividades previstas para a próxima fase são o uso tanto de mais imagens para o treinamento da SVM, quanto de índices estatísticos para reduzir o número de variáveis na classificação o que diminuirá o tempo de execução da metodologia e tornará mais precisa uma extração da região da boca.

AGRADECIMENTOS

Agradecemos à UFMA e ao CNPq, que financiaram e apoiaram esse projeto de pesquisa.

REFERÊNCIAS

BARALDI, A.; PARMIGGIANI, F. An investigation of the textural characteristics

associated with gray level co-occurrence matrix statistical parameters. IEEE Transactions on Geoscience and Remote Sensing, v. 33, n. 2, p.293-304, 1995.

BARTLETT, M. S. et al. Machine learning methods for fully automatic recognition of facial expressions and facial actions. Proceedings of the IEEE Conference on Systems, man & Cybernetics. 2004.

CHANG, C.-C.; LIN, C.-J. *LIBSVM*: a library for support vector machines. 2011. Disponível em: <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>. Acesso em: 20 out. 2011.

CASTRO, F. C. C. *Reconhecimento e localização de padrões em imagens utilizando redes neurais artificiais como estimador de correlação espectral.* 1995. Tese (Mestrado em Engenharia Elétrica) - Programa de Pós-Graduação em Engenharia, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1995.

CORTES, Corinna; VAPNIK, Vladimir N. Support-vector networks, *Machine Learnin*g. n. 20, 1995. Disponível em: <http://www.informatik.unitrier.de/~ley/db/journals/ml/ml20.html#CortesV95>. Acesso em: 20 out. 2011.

CRISTIANINI, N.; SHAWE-TAYLOR, J. *An introduction to support vector machines and other kernel-based learning methods.* Cambridge: University Press, 2000.

HARALICK, R. M.; SHANMUGAM, K. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics,* v. 3, SMC-3, n. 6, p. 610-621, nov. 1973.

PEDRINI, Hélio; SCHWARTZ, William Robson. *Análise de imagens digitais*: princípios, algoritmos, e aplicações. São Paulo: Thomson Learning, 2008.

PHAM, D. T.; ALCOCK, R. J. Automated grading and defect detection: a review. *Forest Products Journal*, v. 48, n. 4, p. 34-42, 1998.

RAFAEL C. et al. *Digital image processing*, 2.ed., Upper Saddle River, NJ: Prentice Hall, 2002.

MITCHELL, T. *Machine learning.* New York: McGraw-Hill. 1997.

VAPNIK, V.N. Statistical learning theory. New York: Wiley, 1998.