

**UNIVERSIDADE FEDERAL DO AMAZONAS  
FACULDADE DE TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

**AVALIAÇÃO DE CLASSIFICADORES HAAR PROJETADOS  
PARA DETECÇÃO DE FACES**

**RAFAEL PADILLA**

**MANAUS  
2012**

**UNIVERSIDADE FEDERAL DO AMAZONAS**  
**FACULDADE DE TECNOLOGIA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

**RAFAEL PADILLA**

**AVALIAÇÃO DE CLASSIFICADORES HAAR PROJETADOS  
PARA DETECÇÃO DE FACES**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade do Amazonas, como requisito parcial para a obtenção do título de Mestre em Engenharia Elétrica, área de concentração Controle e Automação de Sistemas.

Orientador: Prof. Dr. Cícero Ferreira Fernandes Costa Filho  
Co-orientadora: Profa. Dra. Marly Guimarães Fernandes Costa

**MANAUS**  
**2012**

Aos meus exemplos de vida,  
meus pais, irmãos, avós e  
Omar Ricardo Chehayeb que  
sempre me estimularam a dar  
grandes passos. Obrigado por  
serem minha fonte de  
inspiração, apoio e ensino  
diário.

**RAFAEL PADILLA**

**AVALIAÇÃO DE CLASSIFICADORES HAAR PROJETADOS PARA  
DETECÇÃO DE FACES**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica na área de concentração Controle e Automação de Sistemas.

Aprovado em 24 de Setembro de 2012.

**BANCA EXAMINADORA**

Prof. Dr. Cícero Ferreira Fernandes Costa Filho

Universidade Federal do Amazonas- UFAM

Marly Guimarães Fernandes Costa

Universidade Federal do Amazonas- UFAM

Prof. Dr. José Raimundo Gomes Pereira

Universidade Federal do Amazonas- UFAM

Prof. Dr. Jozias Parente de Oliveira

Universidade do Estado do Amazonas- UEA

## **AGRADECIMENTOS**

Agradeço à força maior, que guia e protege minha vida, pelas oportunidades e conquistas. Por me dar força e coragem para enfrentar os momentos difíceis, pelas pessoas queridas, amigos e minha família;

Aos meus orientadores, Prof. Dr. Cícero Ferreira Fernandes Costa Filho e Profa. Dra. Marly Guimarães Fernandes Costa que sempre estiveram me acompanhando, tal qual, pais que ensinam os filhos;

À todos meus familiares e amigos pelo apoio. Roberto, Suzan, Rodrigo e Vinicius Padilla, Dr. Jorge Felipe Franco de Sá, José Maria Veloso Júnior, Ingrid Chehayeb, Tayná Macedo, Sandra Maria Fodra, Carlo Fodra, Fabiana Padilla e Nara Almeida, vocês sempre estiveram ao meu lado nos momentos que mais precisei;

Ao meu amigo e exemplo de dedicação, esforço e conquistas Omar Ricardo Cheyaheb. Agradeço pelos ensinamentos, conversas e lições que aprendi com você;

Aos meus colegas de turma que sempre me incentivaram, ajudaram e me motivaram no desenvolvimento do trabalho e durante a fase de disciplinas;

Aos professores do curso de pós-graduação em engenharia elétrica pelo aprendizado adquirido;

À Universidade Federal do Amazonas e em especial ao Centro de Tecnologia Eletrônica e da Informação – CETELI - pela concessão de toda infra-estrutura para a realização desse trabalho;

À CAPES pela concessão da bolsa de estudos e apoio financeiro;

Ao Instituto Nokia de Tecnologia – INdT – pelo apoio e incentivo à especialização de seus funcionários.

## RESUMO

Este trabalho de dissertação apresenta o desenvolvimento de um novo método de avaliação de classificadores Haar para localização de faces em imagens, com a finalidade de auxiliar pesquisadores na escolha do melhor classificador dentre os disponíveis atualmente, assim como oferecer uma nova métrica para avaliar futuros classificadores de localização facial. A localização facial, primeira etapa de sistemas biométricos pela face, limita as regiões das faces na imagem de entrada. A localização precisa da face ainda é uma tarefa desafiadora, e possibilita uma segmentação da área da face mais eficaz, aumentando a acurácia no processo de identificação e reconhecimento de indivíduos. O *framework* Viola-Jones tem sido amplamente utilizado por pesquisadores para detecção e localização de objetos e faces. Classificadores de detecção facial que utilizam o método Viola-Jones são compartilhados pela comunidade científica e acadêmica. Porém, pouco é discutido sobre a precisão de tais classificadores. Com a aplicação de 10 classificadores em dois bancos de faces distintas entre si (FEI e Yale), uma análise foi realizada e um novo método de avaliação foi proposto. Neste método uma nova métrica foi desenvolvida, levando em consideração *scores* dados a 22 pontos faciais. Os resultados obtidos mostraram-se mais precisos que a única metodologia de avaliação de classificadores de face Haar presente até o momento.

Palavras-chaves: Banco de faces, detecção facial, pontos faciais, haar *wavelets*, detectores Viola-Jones.

## **ABSTRACT**

This dissertation work presents the development of a new evaluation method for Haar classifiers designed to locate faces in images, in order to help researchers choose the best classifier for their own need, as well as to propose a new metric to evaluate future classifiers designed for face location. Face location, the primary step of the vision-based automated systems, finds the face area in the input image. An accurate location of the face is still a challenging task, but it enables a more efficient segmentation, making the identification and recognition of individuals more accurate. Viola-Jones framework has been widely used by researchers in order to detect the location of faces and objects in a given image. Face detection classifiers, used by Viola-Jones framework, are shared by the scientific and academic community. Nevertheless, few works have studied their accuracy. By applying 10 of these classifiers in two different face databases (FEI and Yale), an analysis was made and a new heuristic was created considering scores given to 22 facial landmarks. The results appear to be more accurate than the ones obtained by the only heuristic proposed by another work, developed so far.

**Keywords:** Face and facial feature detection, Haar wavelets, Human computer interaction, face datasets

## LISTA DE FIGURAS

Figura 1.1: Exemplos de faces do banco de faces AR com diferentes poses e oclusões .....	14
Figura 2.1: Imagens utilizadas em sistemas de reconhecimento biométrico facial.....	17
Figura 2.2: Exemplos de características posicionadas sobre uma janela de amostragem .....	18
Figura 2.3: Imagem (a) e sua imagem integral (b) .....	19
Figura 2.4: Exemplos de imagens integrais em diversos pontos.....	20
Figura 2.5: Processo de detecção de faces Viola-Jones.....	20
Figura 2.6: Exemplo de classificadores obtidos por Viola-Jones: (a) e (b) Características obtidas pelo classificador; (c) e (d) Características aplicadas sobre a imagem.....	22
Figura 2.7: Detecção em cascata .....	23
Figura 2.8: Componentes da biblioteca OpenCV .....	24
Figura 2.9: Conjunto de características retangulares utilizadas no OpenCV .....	25
Figura 2.10: Conjunto de características originais em comparação com o conjunto estendido proposto por Lienhart e Maydt (2002) .....	25
Figura 2.11: Aplicação de classificadores para localização da face, olhos, nariz e boca.....	26
Figura 2.12: Exemplos de face de perfil localizadas por um dos classificadores de Bradley (2011) .....	27
Figura 2.13: Face localizada que atende os dois critérios sugeridos por Castrillón <i>et al</i> (2011) .....	28
Figura 3. 1: Diagrama das etapas realizadas .....	29
Figura 3. 2: Exemplos de imagens do banco de faces Yale .....	32
Figura 3. 3: Exemplos de imagens do banco de faces FEI.....	33
Figura 3. 4: Estrutura de um classificador Viola-Jones em XML.....	34
Figura 3. 5: Detectando face: deslocando a sub-janela por toda a imagem.....	34
Figura 3. 6: Estrutura de um estágio utilizado no classificador localizador de objetos .....	35
Figura 3. 7: Pontos marcados em uma face do banco de imagens Yale.....	37
Figura 3. 8: Pontos marcados em uma face do banco de imagens FEI .....	38
Figura 3. 9: Imagens do banco de faces Yale com 22 pontos faciais marcados manualmente	38
Figura 3. 10: Imagens do banco de faces FEI com 22 pontos faciais marcados manualmente	39
Figura 3. 11: Exemplos de faces localizadas com regiões suprimidas.....	40
Figura 3. 12: Avaliação de um localizador de face pelo critério II .....	42
Figura 4. 1: Taxas de acerto dos classificadores com aplicação do critério I.....	44
Figura 4. 2: Exemplos de imagens da base FEI reprovadas pelo critério I .....	45
Figura 4. 3: Exemplos de imagens da base Yale reprovadas pelo critério I.....	45
Figura 4. 4: Resultados obtidos pelo critério II .....	46
Figura 4. 5: Exemplos de imagens da base FEI reprovadas pelo critério II.....	47
Figura 4. 6: Exemplos de imagens da base Yale reprovadas pelo critério II .....	47



## LISTA DE TABELAS

Tabela 1: Banco de imagens faciais .....	30
Tabela 2: Classificadores utilizados neste trabalho .....	36
Tabela 3: Pontos faciais selecionados para inclusão no critério I e sua respectiva pontuação.	41
Tabela 4: Coordenadas de retângulos retornados após aplicação de classificadores nas imagens.....	43
Tabela 5: Taxas de acerto obtidas pelos critérios I e II. ....	51
Tabela 6: Resultados dos critérios I e II para o banco FEI utilizando frequência de ocorrência .....	51
Tabela 7: Resultados dos critérios I e II para o banco Yale utilizando frequência de ocorrência .....	52
Tabela 8: Resultados da aplicação do teste chi-quadrado para o banco FEI.....	52
Tabela 9: Resultados da aplicação do teste chi-quadrado para o banco Yale .....	53

## **LISTA DE QUADROS**

Quadro 2. 1: Algoritmo Adaboost de aprendizagem segundo Viola e Jones (2001) .....	21
--	----

## LISTA DE SIGLAS

CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CETELI	Centro de Pesquisa e Desenvolvimento em Tecnologia Eletrônica e da Informação
UFAM	Universidade Federal do Amazonas
OpenCV	<i>Open Source Computer Vision</i>
SVM	<i>Support Vector Machine</i>

## SUMÁRIO

<b>CAPÍTULO 1: INTRODUÇÃO</b>	<b>13</b>
1.1 OBJETIVO GERAL	15
1.2 OBJETIVOS ESPECÍFICOS	15
1.3 ORGANIZAÇÃO	15
<b>CAPÍTULO 2: REVISÃO BIBLIOGRÁFICA</b>	<b>17</b>
2.1 INTRODUÇÃO	17
2.2 PESQUISA BIBLIOGRÁFICA	17
2.2.1 Detector de objetos Viola-Jones	18
2.2.2 OpenCV	24
2.2.3 Classificadores Haar projetados para localização de faces	26
2.2.4 Avaliações de Classificadores	27
<b>CAPÍTULO 3: MATERIAIS E MÉTODOS</b>	<b>29</b>
3.1 INTRODUÇÃO	29
3.2 BANCO DE IMAGENS	29
3.2.1 Seleção do Banco de Imagens	32
3.2.2 Banco de Imagens Yale	32
3.2.3 Banco de Imagens FEI	33
3.3 SELECÇÃO DOS CLASSIFICADORES	33
3.4 MARCAÇÃO DE PONTOS FACIAIS	37
3.5 CRITÉRIOS DE AVALIAÇÃO	39
3.5.1 Critério I	39
3.5.2 Critério II	41
3.6 AMBIENTE DE IMPLEMENTAÇÃO	42
<b>CAPÍTULO 4: RESULTADOS</b>	<b>44</b>
4.1 INTRODUÇÃO	44
4.2 RESULTADOS OBTIDOS PELO CRITÉRIO I (PROPOSTA DO AUTOR)	44
4.3 RESULTADOS OBTIDOS PELO CRITÉRIO II (CASTRILLÓN <i>et al</i> , 2011)	46
<b>CAPÍTULO 5: DISCUSSÃO E CONCLUSÕES</b>	<b>49</b>
5.1 INTRODUÇÃO	49
5.2 CARACTERÍSTICAS DO MÉTODO PROPOSTO	49
5.3 DISCUSSÃO DOS RESULTADOS	50
5.4 CONCLUSÕES	53
5.5 TRABALHOS FUTUROS	54
<b>REFERÊNCIAS</b>	<b>55</b>
<b>APÊNDICE</b>	<b>62</b>

## INTRODUÇÃO

Biometria é o estudo das características comportamentais ou físicas de seres vivos. A palavra “biometria”, de origem grega, é derivada das palavras “bio” (vida) e “metrics” (medida). Através da aplicação da biometria é possível reconhecer e identificar indivíduos baseando-se em suas características físicas ou comportamentais.

Sistemas automáticos de identificação são capazes de reconhecer indivíduos através de suas características físicas que podem ser encontradas na face, impressão digital, íris, geometria da mão, odor e DNA. Da mesma forma, características comportamentais também são utilizadas para tais fins, podendo estar presentes na voz, no modo de andar e na assinatura (JAIN *et al*, 2006).

Desde os primórdios, seres humanos já utilizam a face para identificar e reconhecer indivíduos familiares e desconhecidos. Pegadas de animais também serviam de referência para a identificação de espécies (JAIN *et al*, 2010).

Graças aos estudos realizados nos últimos 30 anos na área de reconhecimento de indivíduos pela face (BAGHERIAN *et al*, 2008), sistemas comerciais foram desenvolvidos e são utilizados em diversas aplicações: segurança patrimonial, identificação criminal, controle de acesso, localização de indivíduos, imigração em aeroportos, animações, jogos, entretenimento, etc. Por ser não invasiva e não necessitar da vontade do indivíduo se apresentar ao sistema, a identificação através da face é um tipo de identificação biométrica bastante aceitável no meio comercial. A aplicação de tais sistemas reduz significativamente o tempo de identificação de pessoas em filas, evita fraudes e auxilia na captura de criminosos.

O reconhecimento de faces não é uma tarefa trivial, uma vez que faces formam uma classe de objetos bastante similares e suas características (olhos, boca, nariz, etc) possuem, de maneira geral, a mesma configuração geométrica. Diversas pesquisas nesta área mostram que a precisão dos resultados pode ser comprometida por diversos fatores como a variação de poses e de expressões faciais do indivíduo, iluminação excessiva ou precária, baixa resolução da imagem e oclusão da face por pêlos faciais, cabelo ou objetos como óculos e vestimentas. Tais ruídos presentes nas imagens fazem com que a identificação através da face seja uma tarefa bastante complexa para aplicações encontradas no mundo real.

A Figura 1.1 exemplifica imagens de faces com diferentes expressões faciais e com oclusões, dificultando a identificação do indivíduo em aplicações reais.



**Figura 1.1: Exemplos de faces do banco de faces AR com diferentes poses e oclusões**

**Fonte: Martínez e Benavente (1998)**

É possível encontrar na literatura trabalhos que sugerem a combinação de tipos biométricos diferentes (face+voz) visando a obtenção de melhores resultados quando aplicados em situações reais. Apesar de trabalhos com esta abordagem (multimodal) apresentarem resultados mais precisos, trabalhos que utilizam apenas a imagem da face para a identificação também resultam em taxa de reconhecimento alta, grande eficiência e baixo custo computacional (DHAVAL *et al*, 2009).

Geralmente as imagens de face são pré-processadas antes do início da etapa de identificação. Pode-se citar dentre os possíveis pré-processamentos realizados nas imagens adquiridas: binarização, segmentação, redimensionamento, filtros eliminadores de ruídos e rotação.

A localização facial, primeira etapa de sistemas biométricos através da face, limita a região da face na imagem de entrada. Uma localização precisa da face possibilita uma segmentação do contorno facial mais eficaz, aumentando a acurácia no processo de identificação e reconhecimento de indivíduos. Além de ser a fase principal em sistemas que mapeiam a trajetória de indivíduos.

Diversas abordagens são utilizadas para localizar faces e objetos em imagens. A contribuição do *framework* detector de objetos Viola-Jones possibilitou a localização de faces de forma mais rápida e precisa (VIOLA e JONES, 2002). Implementações deste *framework*, tais como *OpenCV*, redundaram em diversos classificadores que foram desenvolvidos por pesquisadores (BRADLEY, 2011; LIENHART *et al*, 2003a; LIENHART *et al*, 2003b). Pelo fato de ter sido utilizado diferentes bancos de faces durante a fase de treinamento, o desempenho destes classificadores varia muito.

Atualmente existe apenas um trabalho na literatura que faz a avaliação de classificadores de face (CASTRILLÓN *et al*, 2011). Este trabalho utiliza apenas um critério para avaliar se a face foi corretamente localizada.. Por outro lado, o número de bancos de imagens de face que possuem pontos faciais marcados ainda é muito pequeno (ARCA *et al*, 2004; HUA-MING *et al*, 2004; COOTES *et al*, 1998; YUN *et al*, 2009; JEN *et al*, 1996; SOHAIL e BHATTACHARYA, 2006; AMARAPUR e PATIL, 2006; FRANK e CHUANG, 2004). Dentro desse contexto, essa dissertação visa contribuir com o avanço desta área ao propor uma metodologia de avaliação de classificadores de face mais robusta e disponibilizar para as bases de dados de imagens de faces utilizadas os pontos fiduciais acima referidos, além de minuciosa revisão da literatura.

## **1.1 OBJETIVO GERAL**

O foco desse trabalho de dissertação está relacionado com a primeira etapa do processo de identificação de indivíduos através da face, qual seja a localização facial. Objetiva-se o desenvolvimento de um método de avaliação de classificadores que utilizam o método Viola-Jones, projetados para localização da face.

## **1.2 OBJETIVOS ESPECÍFICOS**

- Avaliar a acurácia de classificadores propostos para localização de faces, utilizando bancos de imagens de face comuns;
- Propor uma nova medida para avaliação de métodos de localização de faces, baseada na marcação de pontos fiduciais na face e na atribuição de escores;
- Gerar e disponibilizar bases de dados com imagens de faces contendo os pontos fiduciais anteriormente referidos.

## **1.3 ORGANIZAÇÃO**

Esta dissertação de mestrado está dividida nos seguintes capítulos:

- Introdução (Capítulo 1);
- Revisão Bibliográfica (Capítulo 2);
- Materiais e Métodos (Capítulo 3);
- Resultados (Capítulo 4);

- Discussões e Conclusões (Capítulo 5);
- Referências Bibliográficas;
- Apêndice;

No Capítulo 2 é apresentada uma revisão bibliográfica dos artigos mais relevantes da área de localização facial, tendo como foco apreciar como os autores avaliam a precisão de seus métodos.

O Capítulo 3 aborda os principais bancos de imagens de faces e descreve os bancos que melhor atenderam aos requisitos desejáveis para este trabalho. Também serão descritos os classificadores de face que foram avaliados e os critérios propostos para avaliação dos mesmos.

No Capítulo 4 são exibidos os resultados obtidos com os critérios propostos, sendo feita uma comparação entre eles. Apresentam-se, também, as imagens resultantes das várias etapas do processo de localização facial utilizando os diversos classificadores.

No Capítulo 5 os resultados obtidos são analisados e comparados com os apresentados em outras pesquisas da área. Por fim, conclusões do trabalho são apresentadas e sugerem-se trabalhos futuros.

No Apêndice, encontram-se as cópias dos artigos publicados pelo autor (PADILLA *et al.*, 2012), que foram apresentados em duas Conferências Internacionais: ICDIP 2012: International Conference on Digital Image Processing (Veneza, Itália – 11/04/2012 à 13/04/2012), e ICMEDIA Conferência Internacional de Ciências Forenses em Multimídia e Segurança Eletrônica (Brasília, Brasil - 18/09/2012 à 21/09/2012).



---

## REVISÃO BIBLIOGRÁFICA

### 2.1 INTRODUÇÃO

A localização facial, processo de extrema importância para sistemas de reconhecimento biométricos que utilizam a face, tem sido tema de diversas pesquisas. Métodos e métricas avaliando a precisão das faces localizadas por diferentes abordagens foram apresentados em alguns trabalhos.

Neste capítulo abordaremos as principais técnicas, identificadas na literatura, que são aplicadas na etapa de localização facial e métricas utilizadas para avaliar a precisão das mesmas.

A Figura 2.1 mostra diferentes imagens utilizadas em sistemas de reconhecimento biométrico facial. Nota-se que as faces apresentam diferentes expressões faciais, poses e oclusões (com ou sem óculos, olhos abertos ou fechados, vestimentas e pêlos faciais).



Figura 2.1: Imagens utilizadas em sistemas de reconhecimento biométrico facial  
Fonte: Bellhumer et al (1997) e Martinez e Benavente (1998)

### 2.2 PESQUISA BIBLIOGRÁFICA

Nessa dissertação foi realizada uma pesquisa bibliográfica em bases literárias tais como: Springer, MIT Library, IEEEEXplorer, sendo esta última a que forneceu um maior

número de referências. Primeiramente, trabalhos abordando o tema de localização facial através do método Viola-Jones e trabalhos que analisam a localização de pontos fiduciais foram identificados e posteriormente filtrados por relevância. A partir da análise destes artigos, apresentaremos a seguir os mais relevantes dentre eles.

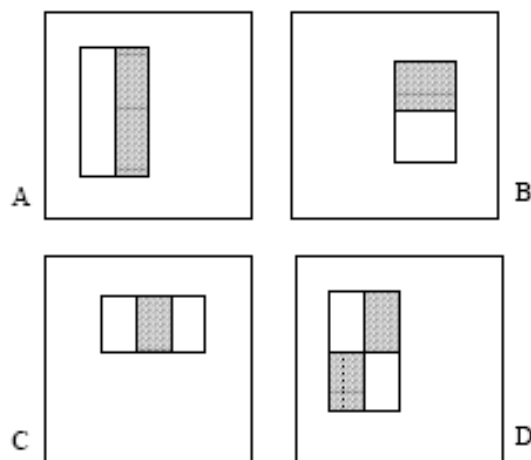
### 2.2.1 Detector de objetos Viola-Jones

Motivados pelo problema de detecção de faces em uma imagem, Viola e Jones (2001) propuseram um detector de faces em tempo real que é amplamente utilizado não somente para a detecção de faces, como também para localização de diferentes objetos em imagens.

O *framework* desenvolvido por Viola e Jones trouxe três grandes contribuições para a comunidade científica:

- A representação da imagem pelo conceito de “imagens integrais”;
- Um novo algoritmo de aprendizado baseado em *Adaboost*;
- Um novo método de localização da face utilizando os classificadores em cascata;

O detector de objetos é baseado no valor das *rectangle features* (chamadas de “características retangulares” nesta dissertação), pois este sistema de detecção tem um desempenho computacional superior ao desempenho de sistemas que analisam valores únicos de pixels. O conceito de *features* (chamadas de “características” nesta dissertação) provém de funções Haar de Base utilizada por Papageorgiou *et al* (1998). Existem basicamente quatro tipos de características, que são demonstradas na Figura 2.2.



**Figura 2.2: Exemplos de características posicionadas sobre uma janela de amostragem**  
 Fonte: Viola e Jones (2001)

Através da Figura 2.2 nota-se que as características podem ser classificadas em características de 2 retângulos (A e B), de 3 retângulos (C) e de 4 retângulos (D).

Cada característica posicionada sobre uma parte da imagem, chamada de janela de amostragem, expressa um valor, o qual é obtido através da soma dos pixels da região branca subtraídos da soma dos pixels da região escura.

O valor expresso por uma característica retangular pode ser rapidamente calculado através da representação da imagem integral. Uma imagem integral é definida como uma matriz com a mesma dimensão da imagem que irá ser representada. O valor na posição  $(x, y)$  de uma imagem integral é dado pela soma dos valores dos pixels localizados acima à esquerda:

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad (2.2)$$

Onde  $ii(x, y)$  é a imagem integral e  $i(x, y)$  é a imagem original.

Na Figura 2.3 mostramos o exemplo de uma imagem e sua imagem integral.

1	1	1
1	1	1
1	1	1

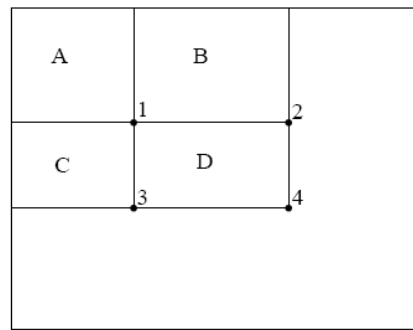
(a)

1	2	3
2	4	6
3	6	9

(b)

**Figura 2.3: Imagem (a) e sua imagem integral (b)**

Para exemplificar o cálculo de uma área qualquer tendo como base uma imagem integral, vejamos a Figura 2.4. A soma dos pixels no retângulo D pode ser computada através de valores obtidos pela soma dos pixels contidos nas áreas A, B, C e D. O valor da imagem integral no ponto 1 é dado pela soma dos pixels contidos no retângulo A. O valor da imagem integral no ponto 2 é dado pela soma dos pixels das áreas A+B, no ponto 3 é dado por A+C e por fim, no ponto 4 é dado pela soma dos valores das áreas A+B+C+D. O valor da imagem integral do ponto 4 também pode ser computado pela soma dos valores dos pontos 4+1, subtraídos da soma dos pontos 2+3 da imagem integral.

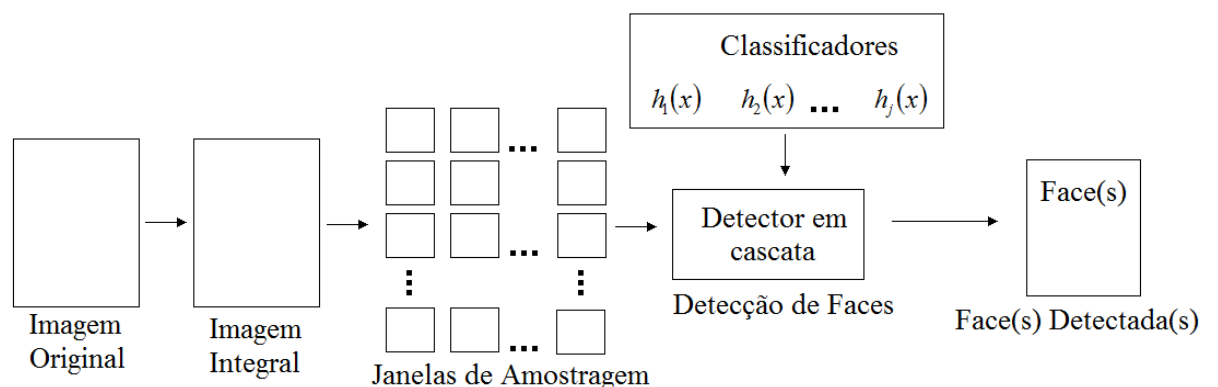


**Figura 2.4: Exemplos de imagens integrais em diversos pontos**  
**Fonte: Viola e Jones (2001)**

Tendo uma imagem qualquer transformada em imagem integral, é possível calcular qualquer característica de uma janela de amostragem desta imagem com grande desempenho computacional, uma vez que para calcular o valor de uma característica formada por 2 retângulos é necessário apenas 6 pontos de referência, oito no caso de uma característica de 3 retângulos e nove para uma característica de 4 retângulos.

Se considerarmos uma imagem de 24x24 pixels, temos um total de 576 pixels, cada um representando uma informação contida na imagem. Nesta mesma imagem é possível termos 134.736 características possíveis utilizando apenas os tipos das características apresentadas na Figura 2.2 (um número bem maior do que a quantidade de pixels).

O detector de faces de Viola-Jones propõe um número bem menor de características combinadas entre si para formar um classificador eficiente. O processo de detecção de faces de Viola-Jones pode ser apresentado conforme a Figura 2.5.



**Figura 2.5: Processo de detecção de faces Viola-Jones**

O processo de detecção de faces inicia-se pela transformação da imagem capturada em imagem integral. Partes da imagem integral, chamadas de janelas de amostragens, são passadas para o detector em cascata que utiliza classificadores treinados para detectar faces. Estes classificadores são formados por características com limiares definidos, obtidos na fase

de treinamento para estabelecer se uma região pertence ou não a uma face. Na fase de treinamento, os parâmetros dos classificadores são ajustados a partir de um algoritmo de treinamento baseado em Adaboost. O Adaboost é um procedimento eficiente porque procura o menor número de características com diferenças significativas entre si. Como entrada para o treinamento é necessário obter imagens positivas (imagens que correspondem a faces) e negativas (imagens que não correspondem a faces). O treinamento combina uma coleção de classificadores fracos para formar um classificador forte. O algoritmo de treinamento está descrito a seguir no quadro abaixo:

**Quadro 2. 1: Algoritmo Adaboost de aprendizagem segundo Viola e Jones (2001)**

1. Obtenha imagens de treinamento  $(x_1, y_1), \dots, (x_n, y_n)$  onde  $y_i = 0, 1$  para imagens negativas e positivas respectivamente

2. Seja  $w_{1,i}$  o vetor de pesos do classificador na primeira interação. Inicialize os pesos  $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$  para  $y_i = 0, 1$  respectivamente, onde  $m$  e  $l$  são os números de imagens positivas e negativas respectivamente.

3. Seja  $t = 1, \dots, T$  o número de interações definido pelo autor. Quanto mais interações forem definidas, mais lento será o processo de treinamento, porém as taxas de acerto das características do classificador serão maiores.

- Normalize os pesos,  $w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^n w_{t,j}}$

sendo  $w_t$  a probabilidade de distribuição.

- Para cada característica,  $j$ , treine um classificador  $h_j$ , o qual é restrito para utilizar uma única característica. O erro é calculado para o respectivo peso  $w_i$  e é dado por  $\epsilon_j = \sum_i w_i |h_j(x_i) - y_i|$

- Escolha o classificador  $h_t$  com o menor erro  $\epsilon_t$

- Atualize os pesos:

$$w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$$

Onde  $e_i = 0$  se a imagem  $x_i$  é classificada corretamente,  $e_i = 1$  caso

contrário e  $\beta_t = \frac{\epsilon_t}{1-\epsilon_t}$ .

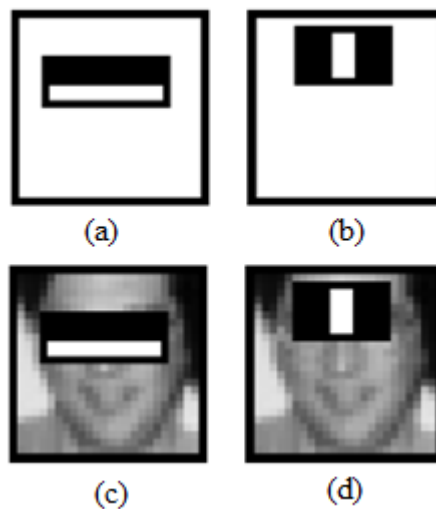
4. O Classificador forte final é dado por:

$$h(x) = \begin{cases} 1 & \sum_{t=1}^T \delta_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \delta_t \\ 0 & \text{caso contrário} \end{cases}$$

onde  $\delta_t = \log \frac{1}{\beta_t}$

A classificação pode ser feita utilizando redes neurais, SVM (*Support Vector Machine*) ou qualquer outra abordagem de aprendizado.

O resultado do treinamento é um conjunto de classificadores compostos por características, sua localização na janela e valores de limiar para cada classificador. Exemplos de classificadores obtidos por Viola-Jones podem ser vistos na Figura 2.6, com dois classificadores, uma característica de 2 retângulos e outra característica de 3 retângulos.

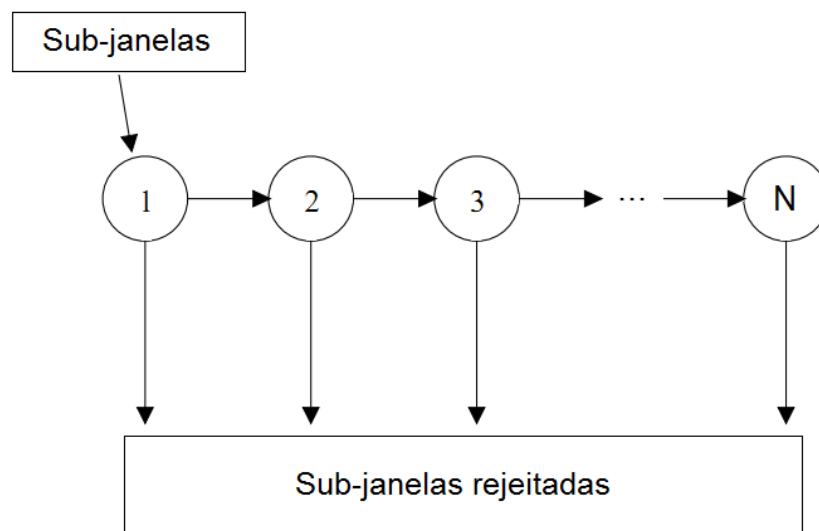


**Figura 2.6:** Exemplo de classificadores obtidos por Viola-Jones: (a) e (b) Características obtidas pelo classificador; (c) e (d) Características aplicadas sobre a imagem.

Fonte: Viola e Jones (2001)

Através da Figura 2.6 (a) é possível perceber que a característica de 2 retângulos representa a diferença de intensidade entre as regiões dos olhos e a parte superior das bochechas. A característica de 3 retângulos, representada pela Figura 2.6 (b), compara a intensidade das regiões dos olhos e a do nariz.

Uma vez obtido o conjunto de características, o processo de teste é feito utilizando uma técnica de verificação em cascata, conforme a Figura 2.7. Cada camada é composta por várias características. Para o classificador comprovar que uma sub-janela possui uma face, esta sub-janela deve passar por todas as N camadas do classificador, de modo que não seja rejeitada por nenhuma delas.



**Figura 2.7: Detecção em cascata**  
**Fonte: Adaptado de Viola e Jones (2001)**

No processo de detecção, janelas de amostragens da imagem são passadas para a primeira camada formada por diversas características. A primeira camada elimina um grande número de amostras negativas. As camadas seguintes, formadas por outros grupos de características, eliminam outras amostras negativas, sendo que os estágios seguintes apresentam um esforço computacional maior que os classificadores anteriores. As janelas de amostragens que não passarem por todas as características de uma camada são descartadas. Após todos os estágios, o número de janelas de amostragens é reduzido drasticamente, restando apenas as janelas de amostragens que contém faces. O classificador em cascata é uma espécie de árvore de decisão.

O experimento de Viola-Jones utilizou 4.916 imagens positivas (faces) e 10.000 imagens negativas (não faces) para treinamento. As imagens positivas e negativas foram obtidas através da internet e recortadas manualmente e redimensionadas para o tamanho de

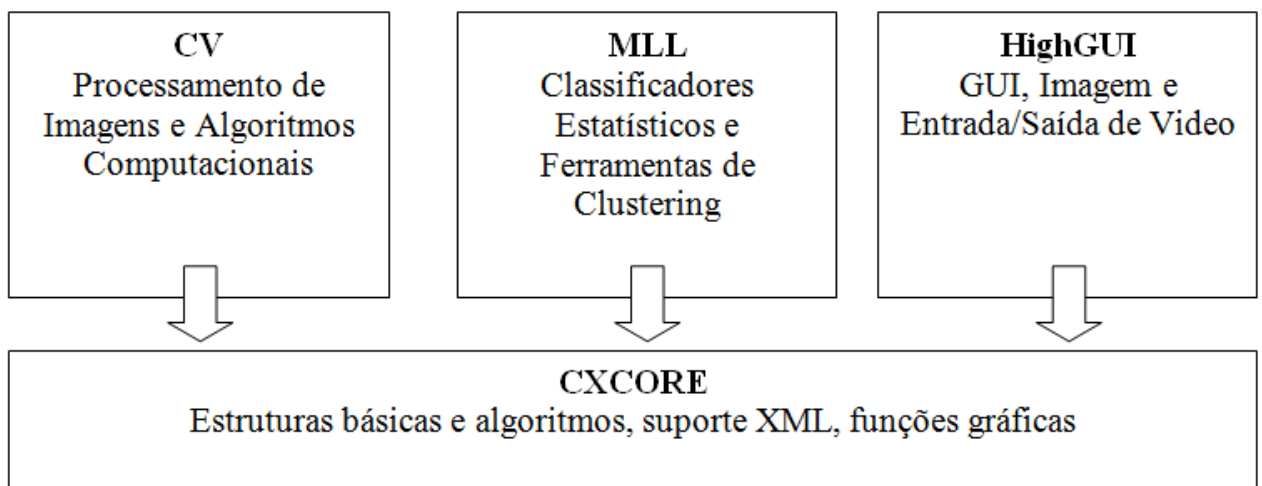
24x24 pixels. O classificador final obtido no experimento dos autores é composto por 32 camadas de classificadores formados por 4.297 características.

### 2.2.2 OpenCV

Após a apresentação do framework de Viola-Jones, diversos autores tomaram a iniciativa de implementá-lo. Dentre as implementações realizadas, o OpenCV® (Open Source Computer Vision) é a mais popular e mais utilizada por pesquisadores (CASTRILLÓN *et al*, 2011).

O OpenCV® é uma biblioteca desenvolvida pela Intel® e atualmente é administrada pela empresa Willow Garage. O OpenCV possui um conjunto de funções desenvolvidas para o processamento de imagens em tempo real. Esta biblioteca, originalmente desenvolvida em linguagem C e C++, pode ser facilmente portada para ser utilizada em diferentes plataformas (C#, Python, Matlab, Ruby e Java). É possível encontrar 500 funções que podem ser aplicadas em projetos de diversas áreas, incluindo engenharia biomédica, segurança, calibração de câmeras e robótica. Esta robusta biblioteca também é formada por funções que envolvem técnicas de reconhecimento de padrões e estatística.

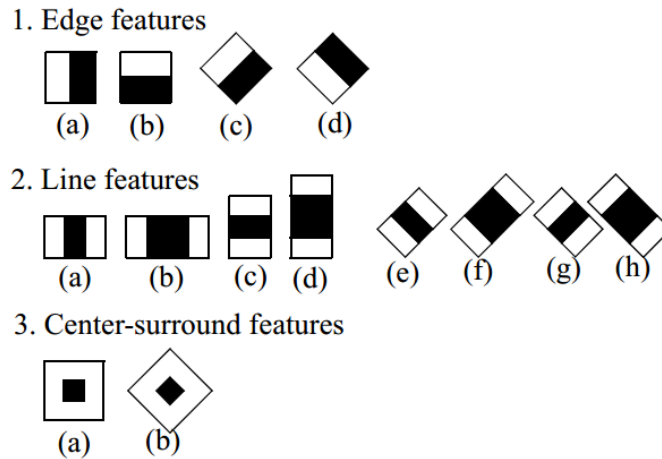
A Figura 2.8 mostra os quatro principais componentes da biblioteca OpenCV.



**Figura 2.8:** Componentes da biblioteca OpenCV

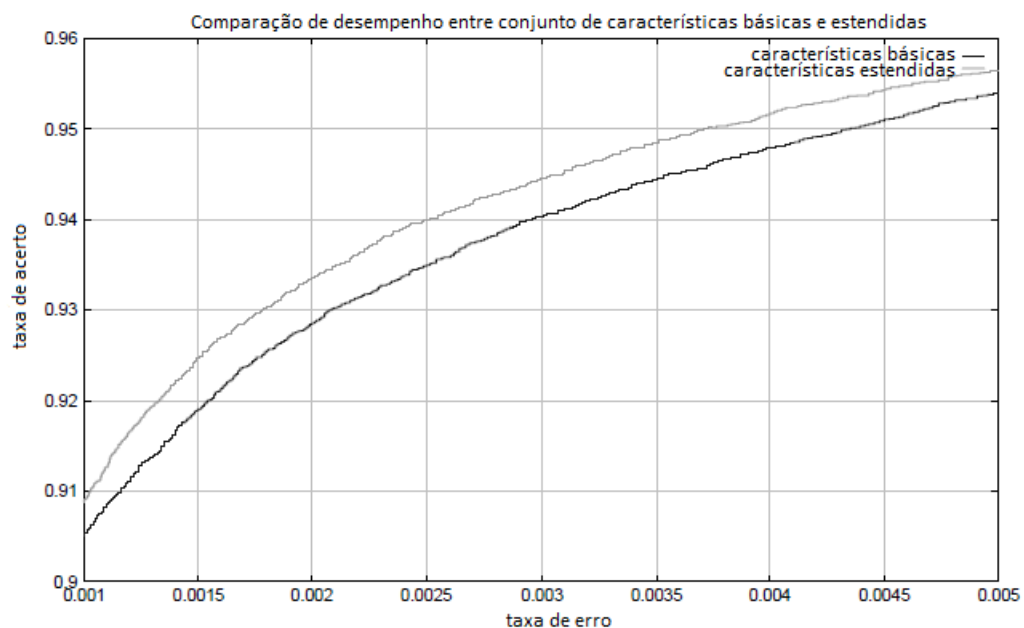
A implementação do *framework* Viola-Jones no OpenCV utiliza o conjunto de características apresentadas por Lienhart e Maydt (2002). Este novo conjunto utiliza as mesmas características retangulares apresentadas por Viola e Jones (2001), porém rotacionadas em 45°. A Figura 2.9 ilustra este novo conjunto.





**Figura 2.9: Conjunto de características retangulares utilizadas no OpenCV**  
**Fonte: Lienhart e Maydt (2002)**

A utilização deste novo conjunto de características diminuiu em 10% a quantidade de alarmes falsos em experimentos realizados com a mesma base de imagens em relação ao conjunto original apresentado por Viola e Jones (2001). A Figura 2.10 mostra duas curvas ROC, uma delas mostrando o desempenho de um sistema de segmentação de faces utilizando as características retangulares originais e outra mostrando o desempenho de um sistema de segmentação de faces utilizando as características retangulares mostradas na Figura 2.9. Como pode ser visto o desempenho do segundo sistema é superior ao desempenho do primeiro.



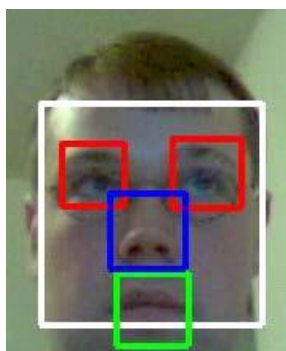
**Figura 2.10: Conjunto de características originais em comparação com o conjunto estendido proposto por Lienhart e Maydt (2002)**  
**Fonte: Lienhart e Maydt (2002)**

### 2.2.3 Classificadores Haar projetados para localização de faces

É possível encontrar na literatura trabalhos que desenvolveram classificadores Haar para localização de faces. (LIENHART *et al*, 2003; WILSON e FERNANDEZ, 2006; BRADLEY, 2011; KRUPPA *et al*, 2003)

Lienhart *et al* (2003a) utilizaram 5000 imagens positivas e 3000 negativas e geraram classificadores capazes de detectar faces de no mínimo 18x18 pixels e 20x20 pixels. Neste trabalho, verificou-se que os de 20x20 pixels apresentam o melhor desempenho na localização de faces frontais. Os autores verificaram que o algoritmo de classificação mais eficiente para separar os conjuntos de imagens positivas e negativas é o Gentle Adaboost.

Em outro trabalho, Lienhart *et al* (2003b) construíram 3 classificadores de face e boca utilizando o método de Viola-Jones com o banco de faces XM2FDB (MESSER *et al*, 1998). Wilson e Fernandez (2006) também criaram classificadores de faces e de regiões faciais (olhos, nariz e boca), porém utilizaram o banco de faces FERET (PHILLIPS *et al*, 2000) para obterem as imagens positivas. A Figura 2.11 ilustra a aplicação do trabalho de Wilson e Fernandez (2006) para localização da face, olhos, nariz e boca em uma face.



**Figura 2.11: Aplicação de classificadores para localização da face, olhos, nariz e boca**  
**Fonte: Wilson e Fernandez (2006)**

Bradley (2011) criou classificadores localizadores de face aptos a detectarem faces de perfil. Utilizou imagens dos bancos FERET (PHILLIPS *et al*, 2000) e CMU (SIM *et al*, 2003) para gerar imagens positivas. A Figura 2.12 mostra exemplos de faces detectadas por um destes classificadores.



**Figura 2.12: Exemplos de face de perfil localizadas por um dos classificadores de Bradley (2011)**  
**Fonte: Bradley (2011)**

## 2.2.4 Avaliações de Classificadores

É possível encontrar na literatura alguns trabalhos que sugerem métricas e métodos de avaliação de métodos localizadores de faces.

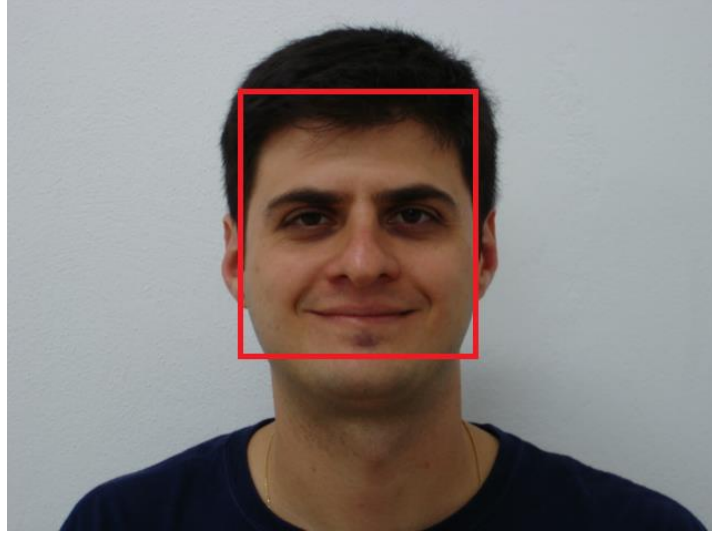
Schmidt *et al* (2007) compararam classificadores para localização de faces e definiram como sendo uma localização precisa da face se a área de intersecção entre o retângulo que representa a face localizada e o retângulo representando a área real da face for maior do que 80% da área dos dois retângulos somados.

Castrillón *et al* (2011) utilizaram dois critérios para avaliar se uma face foi corretamente localizada:

- Todas as características faciais devem estar localizadas na área da face
- A largura da área da face localizada não pode ser maior do que quatro vezes a distância entre os olhos – sugerido por Jesorsky *et al* (2001).

No primeiro critério os autores consideram apenas nariz, boca e olhos como características faciais que devem estar presentes na área da face localizada.

A Figura 2.13 mostra um exemplo onde os dois critérios utilizados por Castrillón *et al* (2011) mostram-se ineficientes, uma vez que, apesar da região facial obedecer aos dois critérios, nota-se que regiões importantes da face (queixo e contorno facial) não estão contidas na região localizada e a região do cabelo (não relevante para sistemas biométricos que utilizam a face) está incluída no retângulo encontrado.



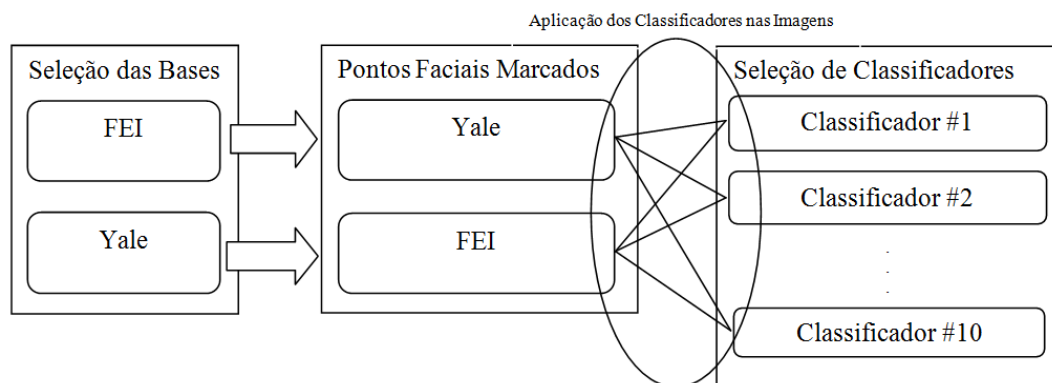
**Figura 2.13:** Face localizada que atende os dois critérios sugeridos por Castrillón *et al* (2011)

## MATERIAIS E MÉTODOS

### 3.1 INTRODUÇÃO

Esse capítulo tem por objetivo principal apresentar o novo método proposto para avaliação de classificadores de face. Como exposto adiante, é proposto uma nova metodologia para marcação dos pontos faciais e os critérios utilizados para pontuação dos classificadores, em função da inclusão ou não desses pontos faciais na face segmentada. Além disso, o capítulo descreve os classificadores avaliados e os bancos de imagem de faces empregados para essa avaliação.

O diagrama em blocos apresentado na Figura 3. 1 ilustra as principais etapas conduzidas neste trabalho para avaliação dos classificadores. Essas etapas serão detalhadas neste capítulo.



**Figura 3. 1: Diagrama das etapas realizadas**

No primeiro bloco da Figura 3.1 selecionamos os bancos de imagens de face utilizados para avaliação dos classificadores. Em seguida, fizemos a marcação dos pontos faciais nas imagens destes bancos. Por último, empregamos os classificadores selecionados e avaliamos os mesmos em função dos critérios propostos nesse trabalho de dissertação. A seguir detalharemos cada uma dessas etapas.

### 3.2 BANCO DE IMAGENS

Dezenas de bancos de imagens de faces estão disponíveis para aplicações de métodos, técnicas e algoritmos voltados ao reconhecimento através da face. Os bancos de imagens faciais diferenciam entre si nos seguintes aspectos:

- Número de imagens por indivíduos
- Ângulo de captura da imagem
- Condições de iluminação
- Expressões faciais
- Resolução da imagem
- Fundo da imagem

A escolha do melhor banco para aplicação de métodos voltados para o reconhecimento de face deve ser feita de acordo com o objetivo de cada trabalho. A Tabela 1 resume alguns bancos de imagens de face, apresentados por Gross (2005).

**Tabela 1: Banco de imagens faciais**  
**Fonte: Adaptado por Gross (2005)**

<b>Banco de Imagens</b>	<b>Resolução das Imagens (pixels)</b>	<b>Número de indivíduos</b>	<b>Número de Imagens</b>	<b>Poses por Indivíduo</b>	<b>Tipos de Iluminação por Indivíduo</b>	<b>Expressões Faciais por Indivíduo</b>	<b>Referência</b>
AR	768 x 576	116	3.288	1	4	4	Martinez e Benavente (1998)
BANCA	720 x 576	208	Indeterminado	1	Indeterminado	1	Bailly-Bailliere <i>et al</i> (2003)
CMU Hyper	640 x 480	54	Indeterminado	1	4	1	Sim <i>et al</i> (2003)
CMU PIE	640 x 486	68	41.368	13	43	3	Sim <i>et al</i> (2003)
FEI	640x480	200	2.800	14	1	2	Thomaz e Giraldi (2010)

**Tabela 1: Banco de imagens faciais**  
**Fonte: Adaptado por Gross (2005)**

<b>Banco de Imagens</b>	<b>Resolução das Imagens (pixels)</b>	<b>Número de indivíduos</b>	<b>Número de Imagens</b>	<b>Poses por Indivíduo</b>	<b>Tipos de Iluminação por Indivíduo</b>	<b>Expressões Faciais por Indivíduo</b>	<b>Referência</b>
FERET	256 x 384	1.199	14.051	9 - 20	2	2	Phillips <i>et al</i> (2000)
Harvard RL	193 x 254	10	Indeterminado	1	77 – 84	1	HALLINAN (1995)
KFDB	640 x 480	1000	52.000	7	16	5	Hwang <i>et al</i> (2003)
ORL	92 x 110	10	400	1	Indeterminado	Variável	Samaria e Harter (1994)
UMIST	220 x 220	20	564	Variável	1	Variável	Graham e Allinson (1998)
U. Texas	720 x 480	284	Indeterminado	Variável	1	Variável	Toole <i>et al</i> (2003)
U. Oulu	428 x 569	125	Indeterminado	1	16	1	Marszalec <i>et al</i> (2000)
XM2VTS	720 x 576	295	Indeterminado	Variável	1	Variável	Messer <i>et al</i> (1999)
Yale	320 x 243	15	165	1	3	6	Bellhumer <i>et al</i> (1997)

### 3.2.1 Seleção do Banco de Imagens

Como o objetivo desse trabalho é avaliar classificadores que utilizam o método proposto por Viola-Jones para localização de faces frontais, escolhemos dois bancos de faces contendo faces frontais: o banco de imagens de Yale (BELLHUMER *et al*, 1997) e o banco de imagens de FEI (THOMAZ *et al*, 2010).

Um dos critérios utilizados para a escolha destes bancos de imagens é a popularidade dos mesmos no meio científico e acadêmico. As imagens do Yale são imagens em nível de cinza enquanto que as do FEI são *true color*. Tais fatos constituíram-se em uma razão adicional pois através destes poder-se-á avaliar como os classificadores de localização de faces se comportam com imagens de colorações diferentes. O tamanho das imagens escolhidas não influencia a localização de faces pelos classificadores, uma vez que a única limitação de cada classificador é o tamanho mínimo das faces que podem ser localizadas. Nesse caso o tamanho das faces é o fato limitante maior de cada classificador.

### 3.2.2 Banco de Imagens Yale

O banco de Yale, que tem este nome por ter sido criado pela Universidade de Yale, foi apresentado por Bellhumer *et al* (1997) e contém imagens faciais de 15 indivíduos, sendo 11 imagens por pessoa. As imagens foram capturadas em escala de cinza sob diferentes condições de iluminação.

Os indivíduos apresentam diferentes expressões faciais (com óculos, tristes, com sono, surpresas e piscando). Cada imagem possui o tamanho de 320x243 pixels.

A Figura 3. 2 mostra algumas imagens do banco de faces Yale:

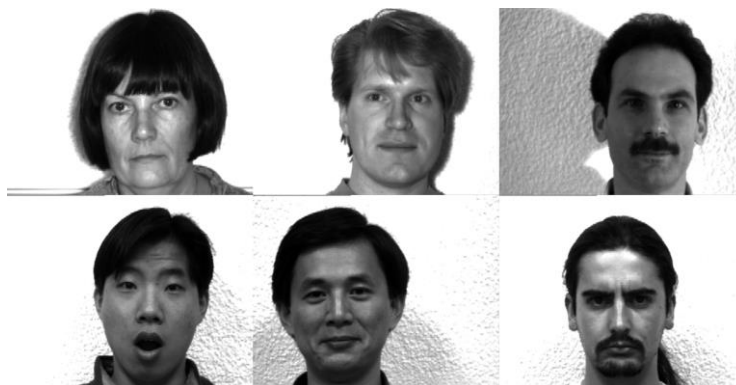


Figura 3. 2: Exemplos de imagens do banco de faces Yale  
Fonte: Bellhumer *et al* (1997)



### 3.2.3 Banco de Imagens FEI

O banco de faces FEI, que tem este nome por ter sido criado pela Fundação Educacional Inaciana Pe. Sabóia de Medeiros, antiga Faculdade de Engenharia Industrial, é o primeiro banco de faces brasileiro. Foi apresentado por Thomaz e Giraldi (2010) e contém 14 imagens para cada um dos 200 indivíduos, totalizando 2.800 imagens.

As imagens são coloridas com diferentes rotações e os indivíduos mostram expressões faciais neutras e sorrindo.

Neste trabalho, utilizamos 2 imagens frontais de cada indivíduo, considerando as imagens com sorriso e expressão neutra, totalizando 400 imagens. O tamanho de cada imagem é 640x480 pixels.

A Figura 3. 3 mostra exemplos de imagens do banco de faces FEI:

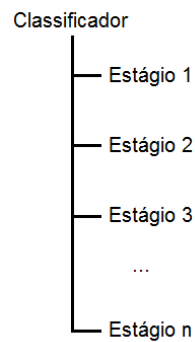


**Figura 3. 3: Exemplos de imagens do banco de faces FEI**  
Fonte: Thomaz e Giraldi (2010)

### 3.3 SELECCÃO DOS CLASSIFICADORES

Graças à implementação no OpenCV das funções de treinamento e geração de classificadores de objetos através do método de Viola-Jones, a comunidade científica foi motivada a desenvolver classificadores localizadores de objetos. Estes classificadores, quando aplicados em uma imagem de entrada é capaz de identificar as regiões (sub-janelas) que contém o objeto alvo a ser localizado.

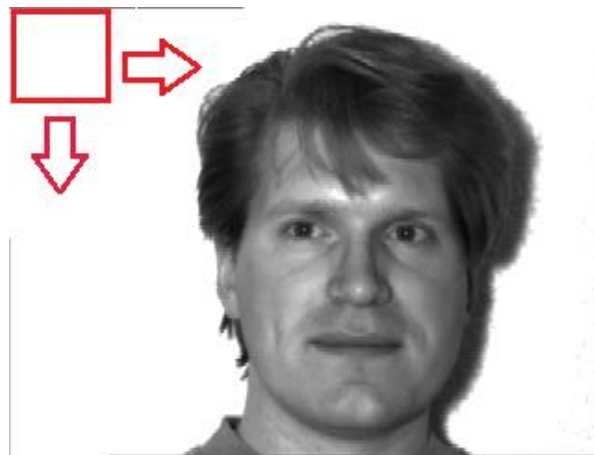
Estes classificadores são arquivos XML (*Extensible Markup Language*) estruturados em estágios como mostra a Figura 3. 4.



**Figura 3. 4: Estrutura de um classificador Viola-Jones em XML**

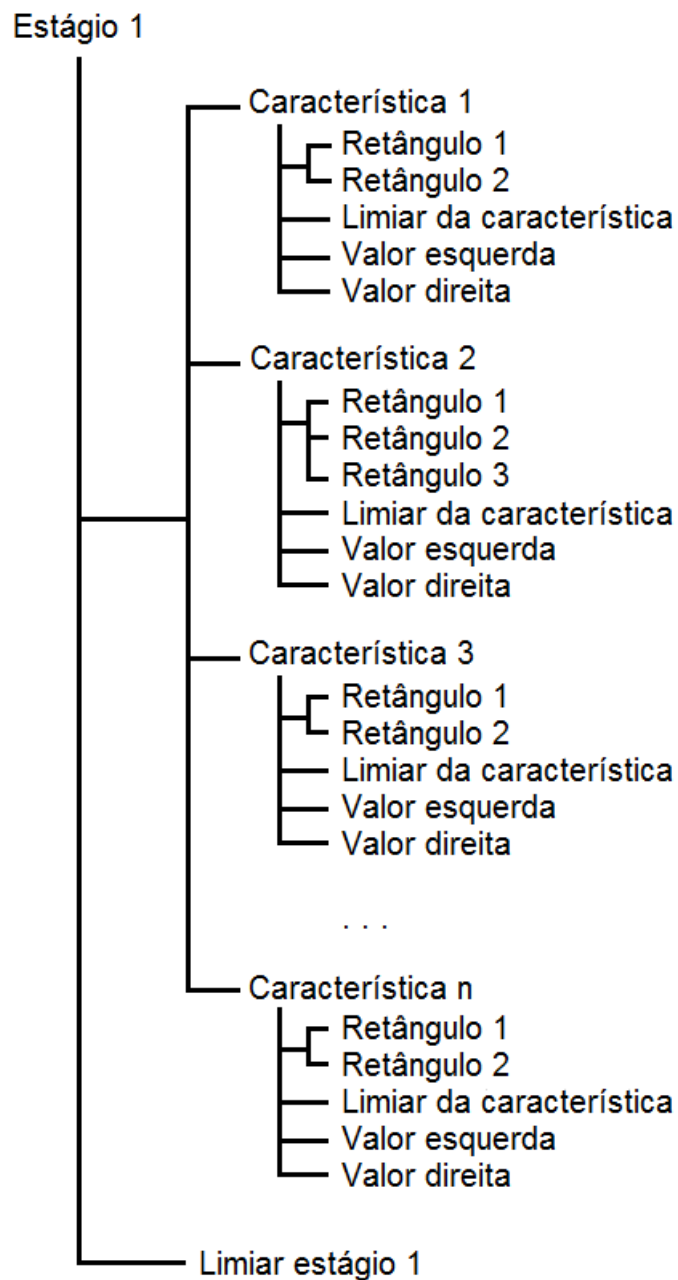
Cada classificador é formado por vários estágios. O classificador determina o tamanho mínimo da sub-janela, que será analisada pelos estágios até que seja reprovada por um deles.

Caso a sub-janela seja reprovada, a mesma é deslocada na imagem, gerando uma nova sub-janela a ser analisada pelo classificador. Ao passar por toda imagem, a sub-janela terá suas dimensões ampliadas e o processo recomeça através de toda a imagem, conforme mostra a Figura 3. 5.



**Figura 3. 5: Detectando face: deslocando a sub-janela por toda aa imagem**

Cada estágio é formado por características e por um limiar de estágio. Cada característica é formada por retângulos, por um limiar de característica, valor a esquerda e valor a direita. A Figura 3. 6 mostra a estrutura de um único estágio.



**Figura 3. 6: Estrutura de um estágio utilizado no classificador localizador de objetos**

Cada característica é formada por 2 ou 3 retângulos, conforme mostrado na Figura 2.9. Cada característica possui um valor a esquerda e um valor a direita. Estes valores são pesos encontrados no processo de treinamento e atribuídos a cada característica, os quais são somados e comparados com o limiar do estágio.

O valor de cada característica é calculado pela aplicação da mesma na sub-janela. Se o valor da característica for menor do que seu limiar, acumula-se o valor da esquerda, caso contrário, o valor acumulado é o valor da direita.

Para uma sub-janela ser aprovada por um estágio, o valor acumulado por todas as características deve ser menor do que o limiar do estágio.

Alguns classificadores localizadores de faces são distribuídos na implementação do OpenCV. Neste trabalho utilizou-se 10 classificadores que são apresentados na Tabela 2.

**Tabela 2: Classificadores utilizados neste trabalho**

<b>Tipo de Classificador</b>	<b>Tamanho</b>	<b>Nº de Estágios</b>	<b>Posição da Face</b>	<b>Referências</b>
FD (haarcascade_frontalface_default.xml)	24x24	25	Frontal	Lienhart <i>et al</i> (2003a); Lienhart <i>et al</i> (2003b)
FA1 (haarcascade_frontalface_alt.xml)	20x20	22	Frontal	Lienhart <i>et al</i> (2003a); Lienhart <i>et al</i> (2003b)
FAT (haarcascade_frontalface_alt_tree.xml)	20x20	46	Frontal	Lienhart <i>et al</i> (2003a); Lienhart <i>et al</i> (2003b)
FA2 (haarcascade_frontalface_alt2.xml)	20x20	20	Frontal	Lienhart <i>et al</i> (2003a); Lienhart <i>et al</i> (2003b)
FW (FaceFrontal.xml)	30x30	19	Frontal	Wilson e Fernandez (2006)
FWQ (FaceQtrTurn.xml)	30x30	20	Rotacionada 1/4	Wilson e Fernandez (2006)
FWH (FaceHlfTurn.xml)	25x30	20	Rotacionada ½	Wilson e Fernandez (2006)
PR (haarcascade_profileface.xml)	20x20	26	Perfil	Bradley (2011)
HS1 (haarcascade_upperbody.xml)	22x18	30	Parte superior do corpo	Kruppa <i>et al</i> (2003)
HS2 (HS.xml)	22x20	19	Cabeça e ombros	Kruppa <i>et al</i> (2003)

A escolha desses classificadores resultou de uma pesquisa bibliográfica criteriosa na literatura, de onde se extraiu as referências listadas na Tabela 2 e seus correspondentes classificadores. Os classificadores FW, FWQ, FWH, HS1 e HS2 foram gentilmente cedidos

pelos autores através de contato realizado. Os classificadores PR, FD, FAT, FA2, FA1 estão disponíveis na versão 2.4 da biblioteca OpenCV.

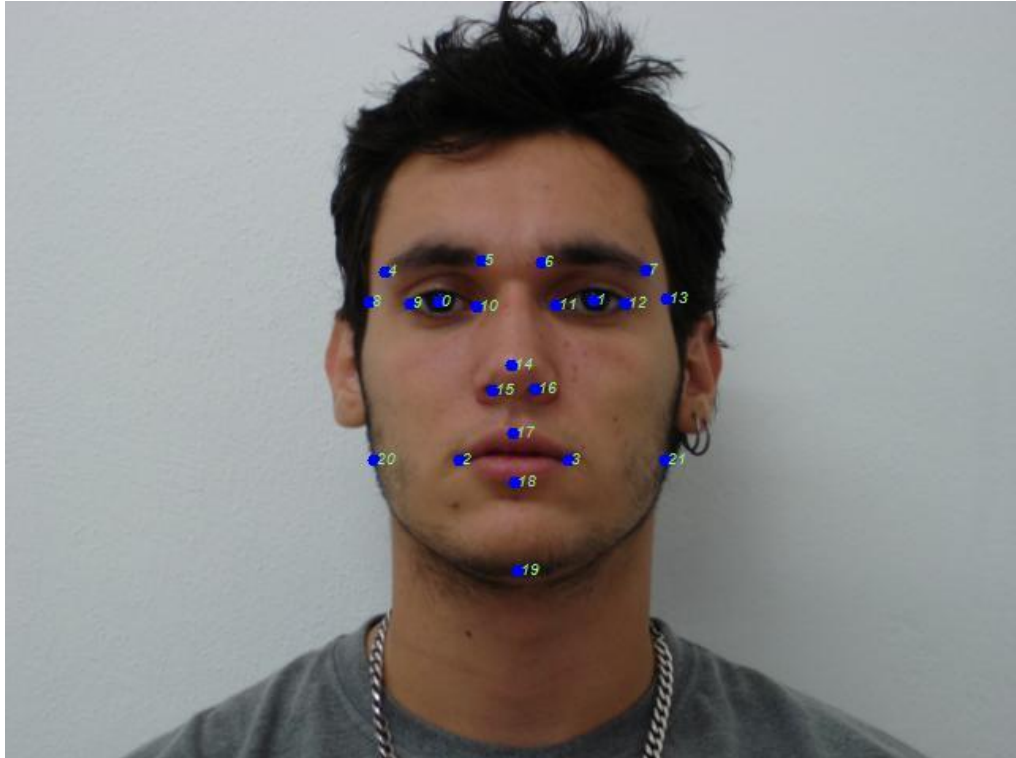
### 3.4 MARCAÇÃO DE PONTOS FACIAIS

A Detecção de pontos faciais é importante não somente para gerar um modelo geométrico da face, mas também para ser usado em detecção de faces. (BEUMER et al, 2002).

O projeto FGnet publicou a localização de 22 pontos faciais de todas as faces do banco de faces AR (MARTINEZ e BENAVENTE, 1998). Neste trabalho nós também marcamos os 22 pontos faciais utilizados no projeto Fgnet nas imagens dos bancos de face Yale e FEI. Esta marcação pode ser vista nas Figuras 3.7 e 3.8.



**Figura 3. 7: Pontos marcados em uma face do banco de imagens Yale**



**Figura 3. 8: Pontos marcados em uma face do banco de imagens FEI**

No total, 565 imagens foram utilizadas. Como será apresentado a seguir, a principal contribuição desta dissertação foi propor um sistema de pontuação para alguns pontos faciais, com o objetivo de avaliação dos classificadores.

O conjunto de pontos marcados pode ser subdividido em dois subconjuntos: pontos externos e pontos internos. O subconjunto  $C_1$  dos pontos externos, que delineam o contorno da face, é formado por:  $C_1 = \{8, 20, 13, 21, 4, 5, 6, 7, 19\}$ . O subconjunto  $C_2$  dos pontos internos é formado por  $C_2 = \{0, 1, 2, 3, 9, 10, 11, 12, 14, 15, 16, 17, 18\}$ .

A Figura 3. 9 mostra outros exemplos de imagens do banco de faces Yale com os 22 pontos faciais marcados manualmente:



**Figura 3. 9: Imagens do banco de faces Yale com 22 pontos faciais marcados manualmente**

A Figura 3. 10 mostra outros exemplos de imagens do banco de faces FEI com os 22 pontos faciais marcados manualmente:



**Figura 3. 10: Imagens do banco de faces FEI com 22 pontos faciais marcados manualmente**

Uma aplicação, utilizando o Microsoft Visual Studio 2010<sup>®</sup>, foi desenvolvida em C# para facilitar a marcação dos pontos faciais, gerando um arquivo .txt no mesmo formato gerado pelo projeto FGNet.

### 3.5 CRITÉRIOS DE AVALIAÇÃO

Nessa dissertação, para avaliação dos classificadores de face, utilizamos dois critérios, que serão explicados a seguir. O critério I, uma proposta nova dentro do escopo dessa dissertação, utiliza os pontos faciais marcados em cada face, conforme descrito nesta seção. O segundo critério foi proposto por Castrillón *et al* (2011).

#### 3.5.1 Critério I

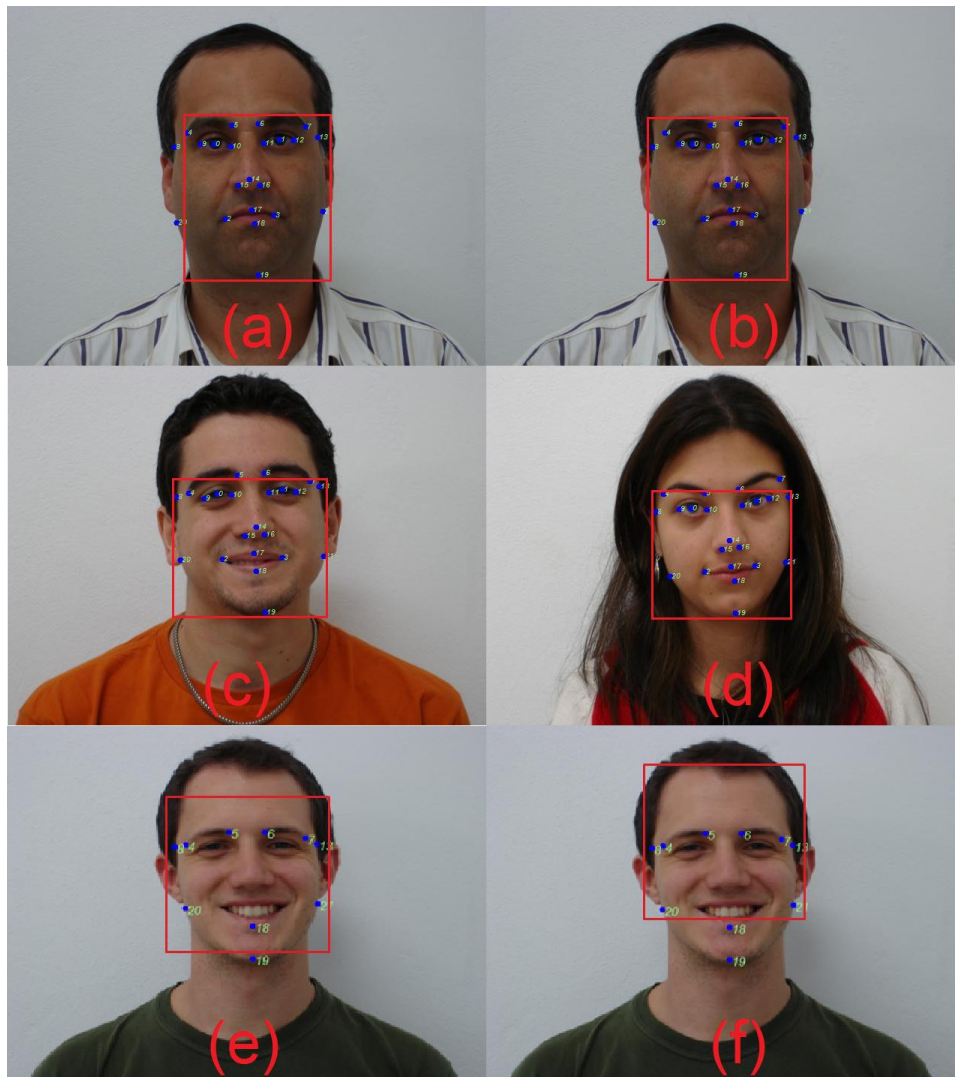
Do ponto de vista da localização da face, considera-se que a mesma fica comprometida quando ocorre uma supressão acentuada de uma dos lados da face, lado esquerdo, lado direito, lado superior e lado inferior. Essa supressão, conforme será mostrado a seguir, pode ser constatada, na maioria dos casos, através da verificação da exclusão de pontos do conjunto  $C_1$ .

Considera-se que uma supressão acentuada do lado esquerdo ocorre quando os pontos de contorno 8 e 20 não são incluídos na região localizada da face. A Figura 3. 11 (a) mostra um exemplo de localização da face onde esses pontos não estão incluídos.

Considera-se que uma supressão acentuada do lado direito ocorre quando os pontos de contorno 13 e 21 não são incluídos na região localizada da face. A Figura 3. 11 (b) mostra um exemplo de localização da face onde esses pontos não estão incluídos.

O contorno superior da face é demarcado através dos pontos 4, 5, 6 e 7. Considera-se que uma supressão acentuada do lado superior ocorre quando dois desses pontos não são incluídos na região localizada da face. A Figura 3. 11 (c) e (d) mostram exemplos de localização da face onde dois desses pontos não estão incluídos.

O contorno inferior da face é demarcado através do ponto 19. Observa-se, no entanto, que uma localização da face excluindo apenas esse ponto, pode não representar uma supressão acentuada da região do queixo, conforme mostrado na Figura 3. 11 (e). Embora a ponta do queixo esteja excluída, a maior parte ainda é preservada na localização. Se, além do ponto 19 for incluído também o ponto 18 pertencente ao conjunto  $C_2$ , verifica-se então que a exclusão desses dois pontos compromete sobremaneira a localização da face, uma vez que, nesse caso, toda a região do queixo é suprimida..



**Figura 3. 11: Exemplos de faces localizadas com regiões suprimidas**



Partindo das considerações feitas nos parágrafos anteriores, construiu-se uma tabela de pontuações, Tabela 3, onde foi atribuído o valor 1 a cada um dos pontos citados acima. Em função da pontuação mostrada nessa tabela, atribuiu-se uma nota à localização feita por um classificador. Assim, se uma localização engloba todos os pontos mostrados na Tabela 3, atribui-se a nota 10 ao classificador. Caso o classificador suprima uma parte considerável de um dos lados da face, conforme conceito expresso anteriormente, o mesmo ganha apenas a nota 8. Assim sendo, considera-se um limiar igual a 8 para afirmar que o classificador não realizou uma boa localização de uma determinada face. Doravante, esse será a definição utilizada para o critério 1 nessa dissertação.

**Tabela 3: Pontos faciais selecionados para inclusão no critério I e sua respectiva pontuação**

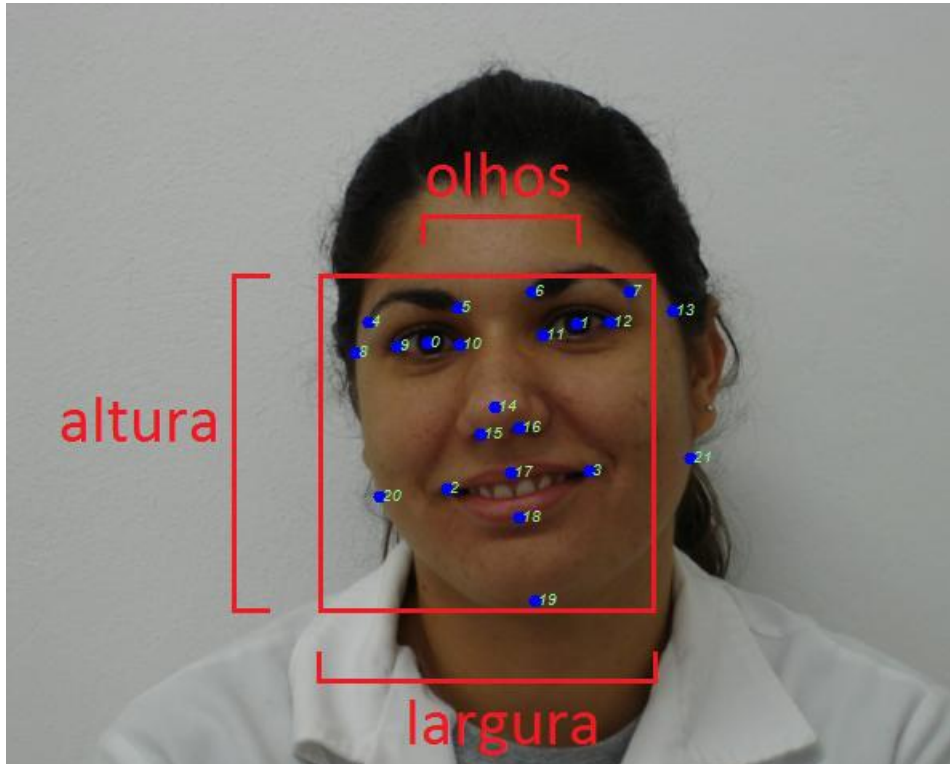
<b>Ponto facial</b>	<b>Descrição</b>	<b>Pontuação</b>
4	Canto direito da sobrancelha direita	1
5	Canto esquerdo da sobrancelha direita	1
6	Canto direito da sobrancelha esquerda	1
7	Canto esquerdo da sobrancelha esquerda	1
8	Limite superior direito da face	1
13	Limite superior esquerdo da face	1
18	Limite inferior labial	1
19	Queixo	1
20	Limite direito da face	1
21	Limite esquerdo da face	1

Verifica-se que a maioria dos pontos pertencentes ao conjunto  $C_2$  não foram utilizados para a determinação do limiar utilizado para o cálculo do critério 1. A marcação dos mesmos, no entanto, evitará trabalhos futuros, em estudos que envolvam a sua determinação, como por exemplo: avaliação de métodos para localização dos olhos, do nariz, etc.

### 3.5.2 Critério II

O critério II é o mesmo critério utilizado por Castrillón *et al* (2011) para avaliar classificadores detectores de face.

Pelo critério II, uma face é corretamente localizada se a largura e altura do retângulo representante da face localizada não são maiores do que quatro vezes a distância entre os olhos.



**Figura 3. 12: Avaliação de um localizador de face pelo critério II**

A distância entre os olhos é calculada pela equação a seguir:

$$D_{\text{direito,esquerdo}} = \sqrt{(y_{\text{esquerdo}} - y_{\text{direito}})^2 + (x_{\text{esquerdo}} - x_{\text{direito}})^2} \quad [1]$$

Onde:

$D_{\text{direito,esquerdo}}$  : Distância entre os olhos

$y_{\text{esquerdo}}$  : Posição  $y$  do olho esquerdo

$y_{\text{direito}}$  : Posição  $y$  do olho direito

$x_{\text{esquerdo}}$  : Posição  $x$  do olho esquerdo

$x_{\text{direito}}$  : Posição  $x$  do olho direito

### 3.6 AMBIENTE DE IMPLEMENTAÇÃO

Os 10 classificadores selecionados e implementado foram aplicados as 565 imagens selecionadas com vistas a que as faces contidas nas imagens fossem localizadas.

Para a aplicação dos classificadores nas imagens de face, utilizou-se a biblioteca do OpenCV no Visual Studio 2010<sup>®</sup>. Um programa foi desenvolvido em C# com vistas a automatizar o cruzamento dos classificadores com as imagens. Um exemplo do resultado do processamento é transcrito na Tabela 3.

**Tabela 4: Coordenadas de retângulos retornados após aplicação de classificadores nas imagens**

<b>Classificador</b>	<b>Imagem</b>	<b>Retângulo da face localizada</b>
FD	subject.01.centerlight.gif	{X=115,Y=67,Width=144,Height=144}
FD	subject01.glasses.gif	{X=101,Y=57,Width=153,Height=153}
FD	subject01.happy.gif	{X=107,Y=61,Width=144,Height=144}
FD	subject01.leftlight.gif	{X=108,Y=65,Width=152,Height=152}
FD	subject01.noglasses.gif	{X=124,Y=64,Width=153,Height=153}
FWQ	subject.01.centerlight.gif	{X=128,Y=96,Width=106,Height=106}
FWQ	subject01.glasses.gif	{X=119,Y=84,Width=115,Height=115}
FWQ	subject01.happy.gif	{X=125,Y=91,Width=107,Height=107}
FWQ	subject01.leftlight.gif	{X=-1,Y=-1,Width=0,Height=0}

Neste exemplo, nota-se que quando o classificador FWQ foi aplicado com a imagem subject01.leftlight.gif, nenhuma face foi encontrada, por isso que o retângulo da face localizada foi identificado com as coordenadas {X=-1, Y=-1, Width=0,Height=0}.

As coordenadas dos pontos faciais coletados das imagens dois bancos de imagens utilizados nessa dissertação resultaram em 565 arquivos .txt. Essa base de dados complementar aos bancos Yale e FEI, gerada nessa dissertação, poderá ser disponibilizada para uso em pesquisas futuras, bastando solicitar através do e-mail eng.rafaelpadilla@gmail.com.

## RESULTADOS

### 4.1 INTRODUÇÃO

Neste capítulo serão apresentados os resultados da aplicação da metodologia de avaliação de classificadores detectores de face proposta, que se constitui de um novo critério proposto pelo autor (critério I) em imagens dos bancos FEI e Yale. Para efeitos de comparação o critério proposto por Castrillón *et al* (2011) (critério II) foi implementado e aplicado às mesmas imagens. Adicionalmente, serão apresentados gráficos comparativos entre os classificadores.

### 4.2 RESULTADOS OBTIDOS PELO CRITÉRIO I (PROPOSTA DO AUTOR)

A Figura 4. 1 mostra as taxas de acerto obtidas obtidos para os 10 classificadores quando o critério I foi aplicado às duas bases.

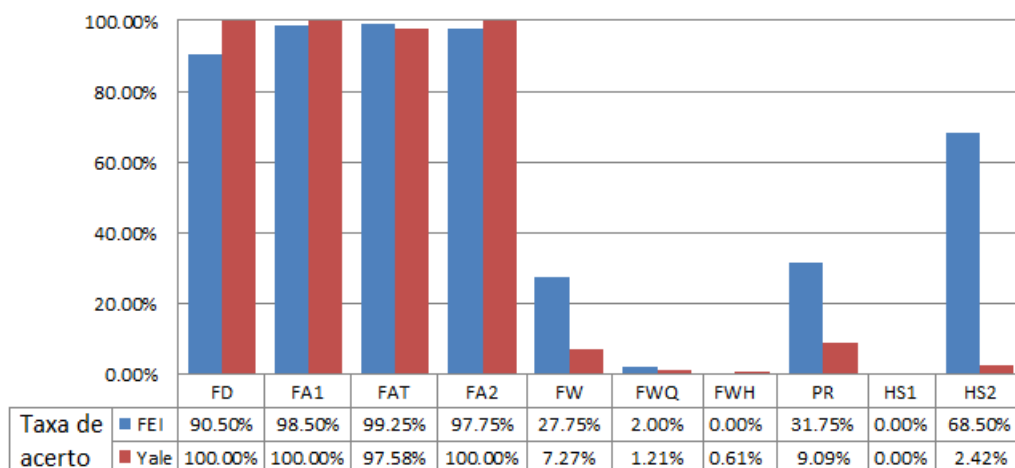


Figura 4. 1: Taxas de acerto dos classificadores com aplicação do critério I

Nas figuras Figura 4. 2 e Figura 4. 3 pode-se ver exemplos de imagens do banco FEI e Yale que foram reprovadas pelo critério I, pois obtiveram pontuações inferiores a 8.

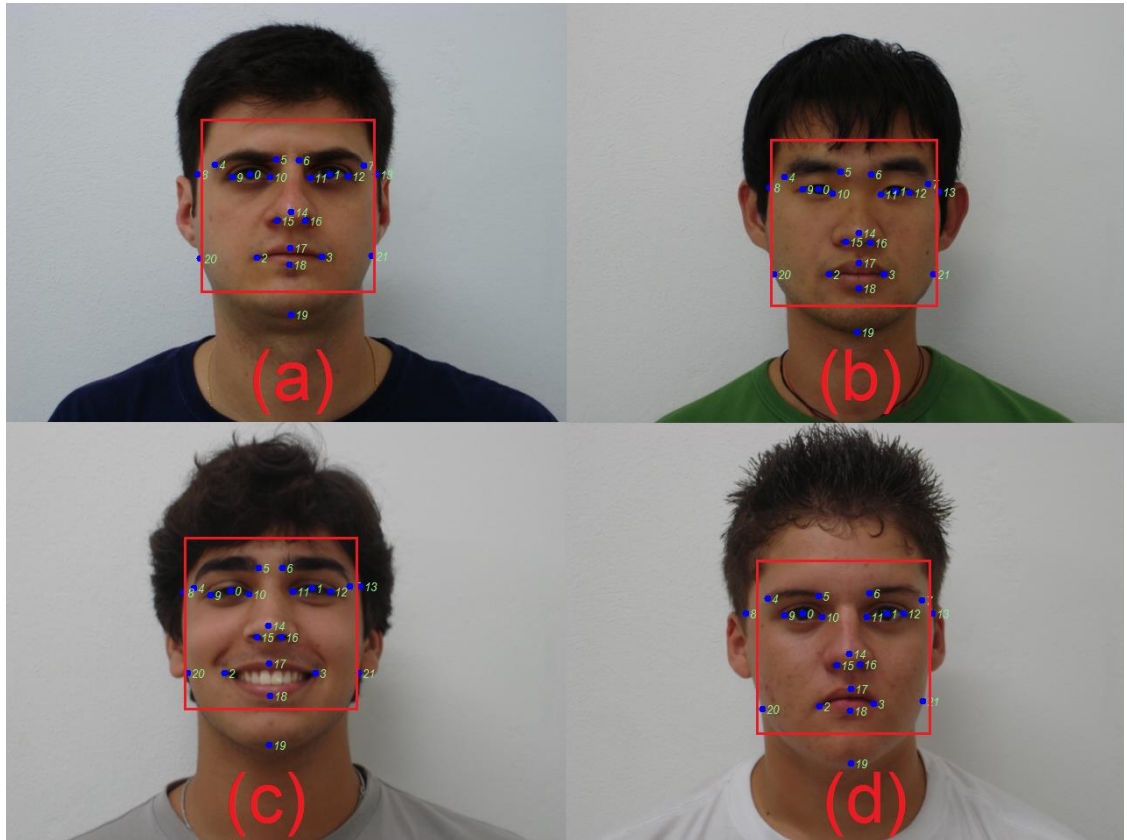


Figura 4. 2: Exemplos de imagens da base FEI reprovadas pelo critério I

Nas figuras Figura 4. 2 (a), (b), (c) e (d) as pontuações obtidas foram 6, 7, 6 e 7 respectivamente.

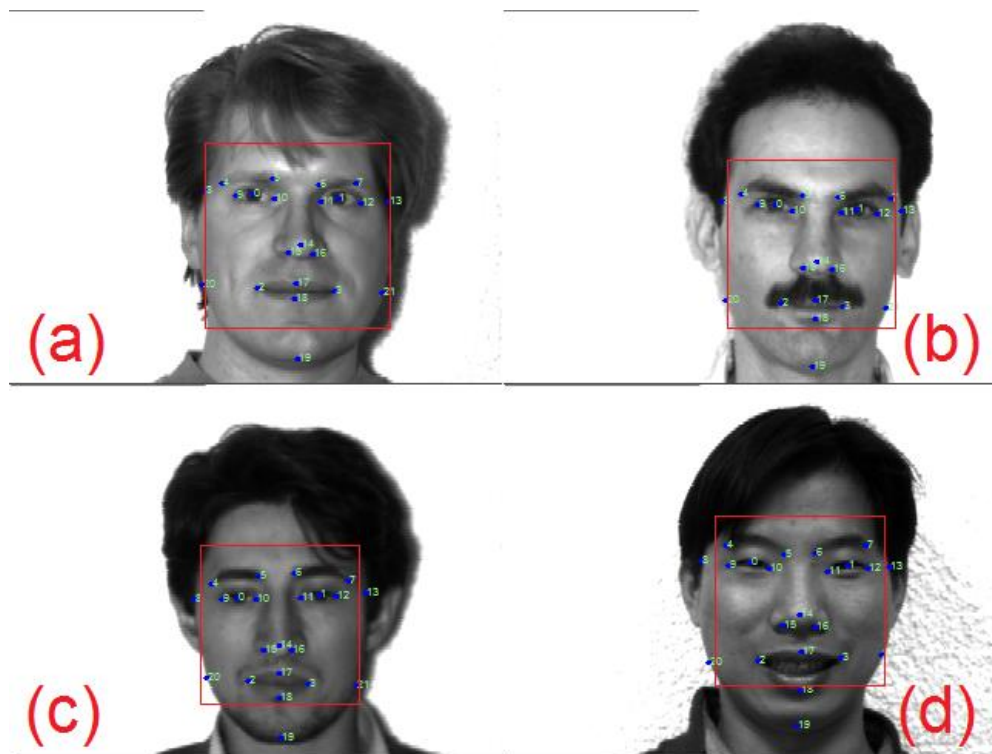


Figura 4. 3: Exemplos de imagens da base Yale reprovadas pelo critério I

Nas figuras Figura 4. 3 (a), (b), (c) e (d) as pontuações obtidas foram 7, 6, 7 e 6 respectivamente.

Em todas elas, nota-se que áreas importantes do contorno da face são suprimidas, principalmente a região do queixo.

### 4.3 RESULTADOS OBTIDOS PELO CRITÉRIO II (CASTRILLÓN *et al*, 2011)

A Figura 4. 4 mostra os resultados obtidos para os 10 classificadores quando o critério II foi utilizado.

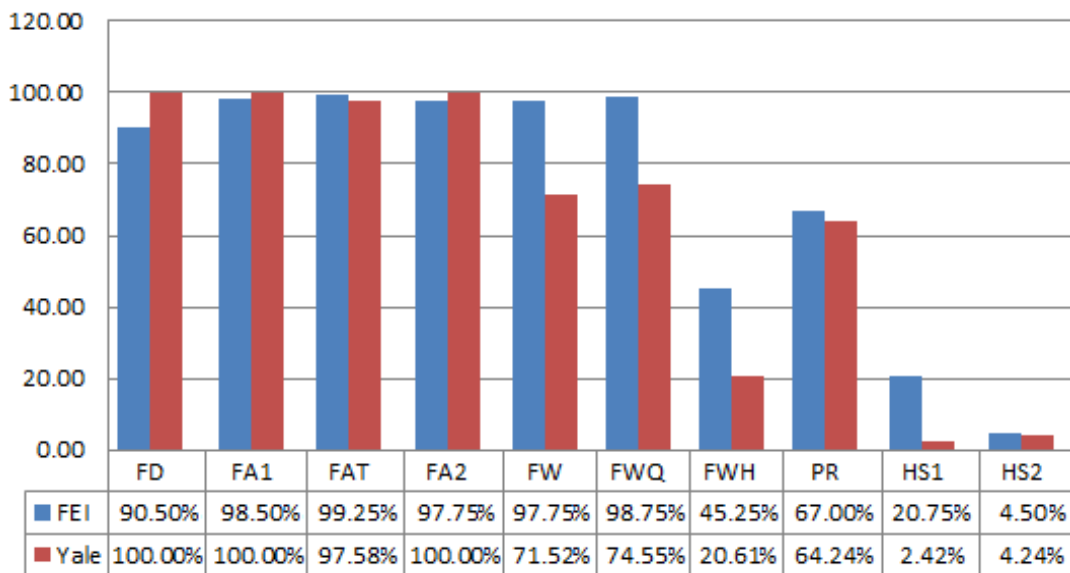


Figura 4. 4: Resultados obtidos pelo critério II

Nas figuras Figura 4. 5 e Figura 4. 6 pode-se ver exemplos de imagens do banco FEI e Yale que foram reprovadas pelo critério II, pois a largura e altura da face localizada são maiores que quatro vezes a distância entre os olhos.

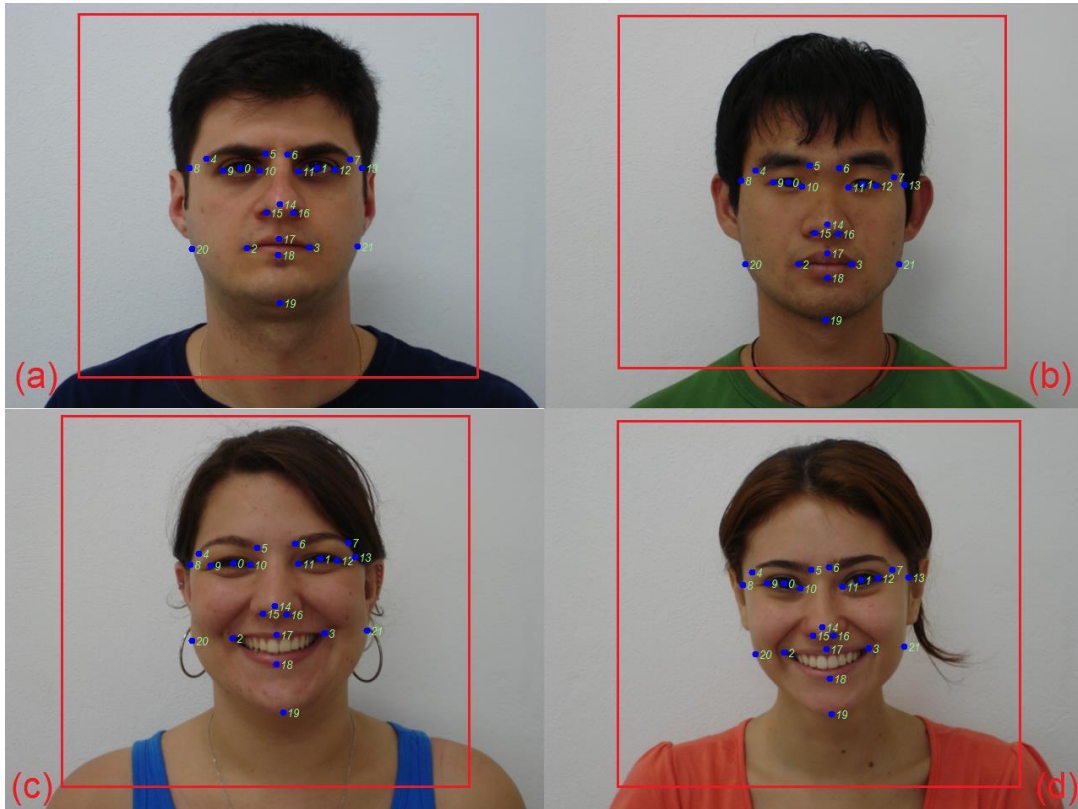


Figura 4. 5: Exemplos de imagens da base FEI reprovadas pelo critério II

As larguras (L), alturas (A) e distância dos olhos (D) das faces da Figura 4. 5 (a), (b), (c) e (d) foram respectivamente: (L=473, A=430 e D=91); (L=458, A=416 e D=88); (L=483, A=439 e D=103,12); (L=475, A=432, D=91).

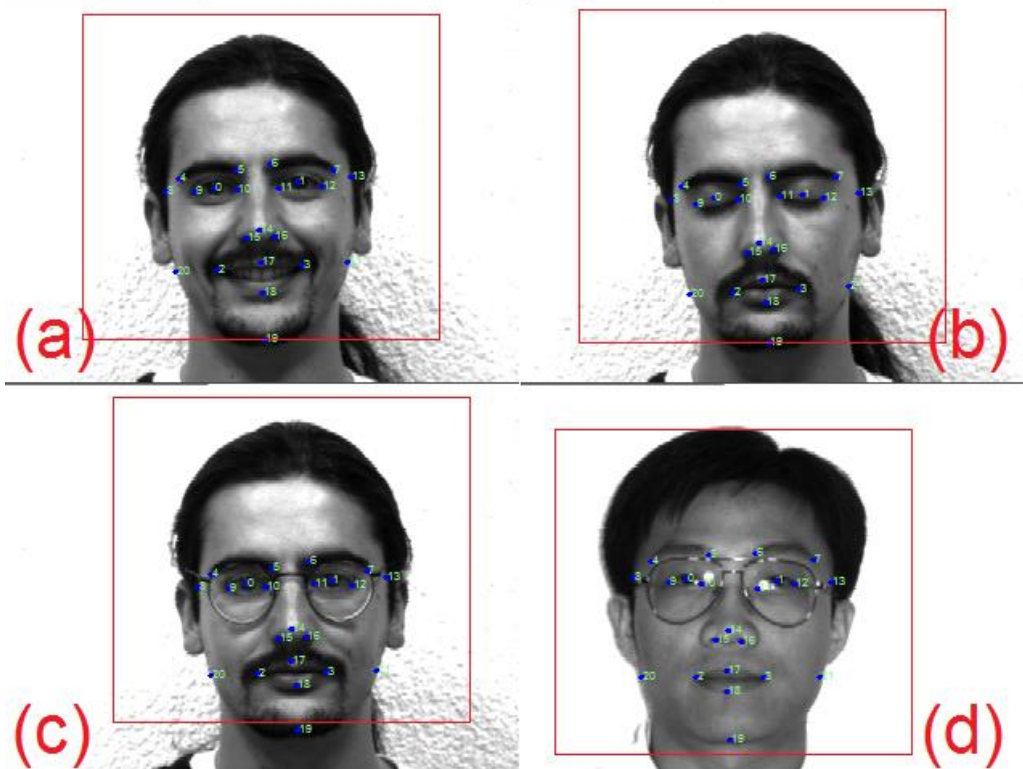


Figura 4. 6: Exemplos de imagens da base Yale reprovadas pelo critério II

As larguras (L), alturas (A) e distância dos olhos (D) das faces da Figura 4. 6 (a), (b), (c) e (d) foram respectivamente: (L=222, A=202 e D=52,15); (L=228, A=207 e D=55,03); (L=222, A=202 e D=53,03); (L=222, A=202, D=51).



## DISCUSSÃO E CONCLUSÕES

### 5.1 INTRODUÇÃO

Finaliza-se essa dissertação com uma discussão do critério proposto para avaliação de classificadores, ressaltando seus aspectos originais e efetuando uma comparação entre a abordagem utilizada pelo autor (critério I) e a abordagem empregada por Castrillón *et al* (2011) (critério II).

Por fim, as limitações do novo critério serão analisadas e os trabalhos futuros sugeridos.

### 5.2 CARACTERÍSTICAS DO MÉTODO PROPOSTO

Como o critério sugerido necessita marcações de pontos faciais, utilizou-se apenas imagens frontais, cujos pontos faciais estão mais presentes. As imagens utilizadas neste trabalho foram selecionadas de dois bancos de imagens diferentes. As imagens do banco FEI utilizadas foram selecionadas de tal forma a se ter somente as imagens de faces frontais deste banco.

Dentre as características observadas nas imagens dos dois bancos, uma que dificultou a marcação manual dos pontos faciais foi o fato desses pontos faciais estarem oclusos pela presença de óculos escuros e cabelos cobrindo regiões da face. Por isso, neste tipo de imagens, estes pontos foram marcados em uma posição estimada.

Com a marcação dos pontos de face, geramos uma base com as coordenadas dos mesmos, possibilitando a obtenção de imagens de certos pontos faciais e regiões (olhos, nariz, boca, etc) através da criação de um programa que extrai estas sub-imagens da face a partir das coordenadas já coletadas.

O critério para avaliação de classificadores de face proposto por esta dissertação apresentou um aspecto fundamental que se reveste de originalidade frente àqueles já existentes na literatura: A necessidade da face localizada apresentar pontos faciais relevantes.

### 5.3 DISCUSSÃO DOS RESULTADOS

É possível encontrar na literatura trabalhos que analisam classificadores de face e de pontos faciais, porém estes trabalhos analisam apenas a velocidade que os classificadores conseguem localizar o suposto objeto (CASTRILLÓN *et al*, 2011; LIENHART *et al*, 2003; SCHMIDT e KASINSKI, 2008).

O único trabalho encontrado que analisa a precisão de classificadores de face é um trabalho de 2011 de Castillón *et al*. Porém este trabalho limita-se apenas em avaliar o tamanho da face encontrada pelos classificadores estudados.

Ao se efetuar a localização da face com os classificadores descritos no capítulo 3, utilizando as imagens dos dois bancos também apresentados no mesmo capítulo, em relação aos resultados obtidos pelo critério I, podemos verificar que:

- Os classificadores FA1, FAT e FA2 obtiveram os melhores resultados.
- Ao se aplicarem os classificadores FA1 e FA2 no banco de imagens Yale, 100% das faces obtiveram pontuação igual ou maior que 8.
- O classificador FAT obteve melhores resultados quando aplicado no banco de imagens FEI, tendo 99,25% de faces aprovadas.
- Os classificadores HS2, HS1, PR, FWH e FWQ não foram gerados para detectarem faces frontais, portanto, uma vez que as imagens utilizadas em ambos são imagens contendo faces frontais, já era de se esperar que apresentassem baixa taxa de aprovação.

Em relação aos resultados obtidos pelo critério II, verificou-se que:

- Os melhores resultados obtidos no banco de Yale foram com os classificadores FD, FA1 e FA2, onde 100% das imagens tiveram suas faces localizadas, com altura e largura menores que quatro vezes a distância entre os olhos.
- Com o banco FEI, os classificadores FAT, FWQ e FA1 os resultados obtidos foram 99,25%, 98,75% e 98,50% respectivamente.
- Notou-se que o classificador FWQ apresentou bons resultados com este critério, porém foi projetado para detectar faces rotacionadas em 45°.

A Tabela 5 agrupa todos os resultados, facilitando a comparação entre os dois classificadores nos dois bancos de imagens utilizados.

**Tabela 5: Taxas de acerto obtidas pelos critérios I e II.**

Classificador	Critério I (proposta do autor)		Critério II (CASTRILLÓN <i>et al</i> , 2011)	
	FEI	Yale	FEI	Yale
FD	90,50%	100,00%	90,50%	100,00%
FA1	98,50%	100,00%	98,50%	100,00%
FAT	99,25%	97,58%	99,25%	97,58%
FA2	97,75%	100,00%	97,75%	100,00%
FW	27,75%	7,27%	97,75%	71,52%
FWQ	2,00%	1,21%	98,75%	74,55%
FWH	0,00%	0,61%	45,25%	20,61%
PR	31,75%	9,09%	67,00%	64,25%
HS1	0,00%	0,00%	20,75%	2,42%
HS2	68,50%	2,42%	4,50%	4,24%

Nas Tabelas 6 e 7 são mostramos, para os bancos FEI e Yale, respectivamente, os resultados semelhantes aos apresentados na Tabela 5, só que utilizando frequências de ocorrência ao invés de percentuais.

**Tabela 6: Resultados dos critérios I e II para o banco FEI utilizando frequência de ocorrência**

Classificador	Critério I		Critério II	
	Acertos	Erros	Acertos	Erros
FD	362	38	362	38
FA1	394	6	394	6
FAT	397	3	397	3
FA2	391	9	391	9
FW	111	289	391	9
FWQ	8	392	395	5
FWH	0	400	181	219
PR	127	273	268	132
HS1	0	400	83	317
HS2	274	126	18	382

**Tabela 7: Resultados dos critérios I e II para o banco Yale utilizando frequência de ocorrência**

Classificador	Critério I		Critério II	
	Acertos	Erros	Acertos	Erros
FD	165	0	165	0
FA1	165	0	165	0
FAT	161	4	161	4
FA2	165	0	165	0
FW	1	164	118	47
FWQ	2	163	123	42
FWH	1	164	34	313
PR	15	150	106	59
HS1	0	165	4	161
HS2	4	161	7	158

Para verificar a significância estatística dos resultados foi aplicado nos dados das Tabelas 6 e 7 o teste de significância estatística chi-quadrado. O número de graus de liberdade nesse caso foi 3 e, para  $p < 0,01$ , o valor crítico foi 11,35.

A Tabela 8 mostra o resultado para o banco FEI, enquanto que a Tabela 9 mostra o resultado para o banco de Yale.

**Tabela 8: Resultados da aplicação do teste chi-quadrado para o banco FEI**

Classificador	Valor da distribuição chi-quadrado ( $\chi^2$ )	Comparação com grau de liberdade
FD	0	$p > 0,01$
FA1	0	$p > 0,01$
FAT	0	$p > 0,01$
FA2	0	$p > 0,01$
FW	419	$p < 0,01$
FWQ	748	$p < 0,01$
FWH	233	$p < 0,01$
PR	99	$p < 0,01$
HS1	92	$p < 0,01$
HS2	353	$p < 0,01$

**Tabela 9: Resultados da aplicação do teste chi-quadrado para o banco Yale**

<b>Classificador</b>	<b>Valor da distribuição chi-quadrado (<math>\chi^2</math>)</b>	<b>Comparação com grau de liberdade</b>
FD	0	$p > 0,01$
FA1	0	$p > 0,01$
FAT	0	$p > 0,01$
FA2	0	$p > 0,01$
FW	179	$p < 0,01$
FWQ	188	$p < 0,01$
FWH	15	$p < 0,01$
PR	108	$p < 0,01$
HS1	4	$p > 0,01$
HS2	0,8	$p > 0,01$

## 5.4 CONCLUSÕES

A partir das Tabelas 5, 8 e 9, algumas conclusões podem ser tiradas.

1. Os melhores índices de aprovação foram alcançados pelos classificadores FA1, FA2 e FAT.
2. Nota-se que o índice de 100% de aprovação só foi conseguido em alguns classificadores, utilizando-se a base de imagens Yale;
3. Os classificadores FD, FA1, FAT, FA2 tiveram índices de aprovação semelhantes ao se utilizar os critérios I e II, em ambas as bases. A avaliação dos mesmos através dos critérios I e II não se mostrou estatisticamente significativa ao nível de 1%;
4. Os classificadores FW, FWQ, FWH, PR e HS1 tiveram índices de aprovação bem menores, em ambas as bases, ao se utilizar o critério I do que o critério II. Na base FEI as diferenças observadas entre os critérios I e II foram estatisticamente significante ao nível de 1%. Na base Yale os critérios FW, FWQ e FWH apresentaram diferenças estatisticamente significantes, ao nível e 1%, aos serem avaliados pelos critérios I e II. Porém, os critérios PR e HS1 não apresentaram diferenças estatisticamente significantes, ao nível de 1%, aos serem avaliados pelos critérios I e II;
5. O classificador HS2 teve um índice aprovação maior na base FEI ao se utilizar o critério I, enquanto que o mesmo teve um índice de aprovação maior na base Yale ao se utilizar o critério II.

Assim sendo, a partir das conclusões anteriores podemos afirmar que:

- Os melhores classificadores avaliados foram FA1, FA2 e FAT.
- Em termos dos percentuais de aprovação, o critério I, proposto nessa dissertação, foi, na maioria dos casos, mais exigente que o critério II. A única exceção a essa regra foi verificada com o classificador HS2, quando avaliado na base FEI.

Embora o critério proposto seja bastante rigoroso quanto à presença de pontos faciais na imagem de face obtida, não limitamos o tamanho desta região. Por isto, a aplicação de um critério adicional que avalia o tamanho da área facial faz-se necessário.

A combinação do critério proposto nesta dissertação com o critério de Castrillón *et al* (2011), que avalia o tamanho da área facial, possibilita uma melhor avaliação dos classificadores.

O método proposto limita-se apenas em avaliar classificadores aplicados em imagens de faces frontais. Imagens com rotações superiores à 45° graus não deverão ser abordadas com o critério sugerido, pois alguns pontos faciais necessários para a marcação não estarão presentes na imagem.

## **5.5 TRABALHOS FUTUROS**

Como trabalhos futuros propõem-se a criação de um classificador de faces, obtido através do método Viola-Jones, cujas imagens utilizadas no treinamento do classificador sejam de diversos bancos de imagens faciais, para que tais imagens possuam diferentes posições faciais, expressões, e adquiridas em ambientes com grande variedade de iluminação. Desta forma, o classificador resultante será robusto suficiente para detectar imagens faciais em diferentes condições e expressões. Propõe-se ainda que o mesmo seja avaliado através do critério de avaliação proposto nessa dissertação.

## REFERÊNCIAS

---

AMARAPUR, B.; PATIL, N. "The Facial Features Extraction for Face Recognition Based on Geometrical Approach", 2006 Canadian Conference on Electrical and Computer Engineering, 2006: 1936-1939.

ARCA, S.; CAMPADELLI, P.; LANZAROTTI, R. An Efficient Method to Detect Facial Fiducial Points for Face Recognition. *17th International Conference on Pattern Recognition*, 2004, Washington DC, Vol. 1, 2004: p. 532-535.

BAGHERIAN, E.; RAHMAT, R. "Facial feature extraction for face recognition: a review" *International Symposium Information Technology*, 2008. Vol. 2, Agosto 2008: 1-9.

BAILLY-BAILLIERE, S; BENGIO F.; BIMBOT, M; HAMOUZ M.; KITTLER, J.; MATAS, K.; MESSER, V.; POPOVICI, F.; POREE, "The BANCA Database and Evaluation Protocol." *Audio- and Video-Based Biometric Person Authentication*, 2003: 625–638.

BELLHUMER P. N.; HESPANHA J.; KRIEGMAN, D. "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection" *IEEE Transactions on Pattern Analysis and Machine Intelligence, Special Issue on Face Recognition*, 1997: 711-720.

BEUMER, M. G.; TAO, Q.; BAZEN, M.; VELDHUIS, N. J. "A Landmark Paper in Face Recognition" *IEEE International Conference on Automatic Face and Gesture Recognition*, 2002: 73-78.

BRADLEY, D. "Profile face detection" Disponível em: <<http://www.cs.cmu.edu/~dbradley/publications/bradley-iurac-03.swf>> Acesso em 10 dez. 2011.

CASTRILLÓN, M.; DÉNIZ, O.; HERNÁNDEZ, D.; LORENZO, J. "A Comparison of Face and Facial Feature Detectors based on the Viola-Jones General Object Detection Framework" *Machine Vision and Applications*, 2011 Vol. 22.

COOTES, T.F.; EDWARDS, G. J.; TAYLOR, C. J. "Active Appearance Models", *European Conference on Computer Vision*, 1998, Vol. 2, 1998:484-498.

DHAVAL, S.; KYU J. H.; SHRIKANTH, S. N. "A Low-Complexity Dynamic Face-Voice Feature Fusion Approach to Multimodal Person Recognition" *IEEE International Symposium on Multimedia*, 2009: 24-31.

ETEMAD, K.; CHELLAPPA, R. "Discriminant Analysis for Recognition of Human Face Images" *Journal of the Optical Society of America*, 1997. Vol. 14, No. 8, 1997: 1724-1733.

FAGERTUN, J, "Face Recognition", *Technical University of Denmark (DTU)*, Dinamarca, 2005.

FRANK, Y. S.; CHUANG, C. "Automatic Extraction of Head and Face Boundaries and Facial Features" *Informatics and Computer Science: An International Journal*, 2004: 117-130.

GANG, Y.; JIAWEI, L.; JIAYY, L.; QINGLI, M.; MING, Y, "Illumination Variation in Face Recognition: A Review" *IEEE Second International Conference on Intelligent Networks and Intelligent Systems*, 2009: 309-311.

GEORGHIADES, A.; KRIEGMAN, D.; BELLHUMER, P. "From Few to Many: Generative Models for Recognition under Variable Pose and Illumination" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001: 643–660.

GOMATHI, E.; BASKARAN, K. "Recognition of Faces Using Improved Principal Component Analysis" *International Conference on Machine Learning and Computing*, 2010: 198-201.



GRAHAM, D. B.; ALLINSON, N. M. "Characterizing Virtual Eigensignatures for General Purpose Face Recognition." *Face Recognition: From Theory to Applications*, NATO ASI Series F, Computer and Systems Sciences, Vol. 163, 1998: 446-456.

GROSS, R. "Face Databases." *Handbook of Face Recognition*. Springer, Heidelberg, 2005: 301-327.

HALLINAN, P.; "A Deformable Model for Face Recognition under Arbitrary Lighting Conditions", Harvard University, 1995.

HEISELE, B.; HO, P.; POGGIO, T. "Face Recognition with Support Vector Machines: Global versus Component-based Approach" *IEEE International Conference on Computer Vision*, 2001. Vol. 2, 2011: 688-694.

HUA-MING, L. MING-QUAN, Z.; GUO-HUA, G. "Rapid Pose Estimation of Mongolian Faces Using Projective Geometry", 2004, *Proceedings. International Symposium on Information Theory*, 2004: 171- 176.

HWANG, B. W.; BYUN, H.; ROH,, M. C.; LEE, S.W. "Performance Evaluation of Face Recognition Algorithms on the Asian face database". *Audio- and Video-Based Biometric Person Authentication*, 2003: 557–565.

JAIN, A. K.; ROSS, A.; PANKANTI, S. "Biometrics: A Tool for Information Security", *IEEE Transactions on Information Forensics and Security*, Junho 2006, Vol. 1, No. 2, 2006: 125-143.

JAIN, A. K.; KUMAR, A. "Biometrics of Next Generation: An Overview", *Second Generation Biometrics*, Springer, 2010.

JENG, S.; LIAO, H.; LIU, Y.; CHERN, M. "An efficient approach for facial feature detection using geometrical face model" *Proceedings of the 13th International Conference on Pattern Recognition*, 1996, Vol.3, 1996: 426- 430.

JESORSKY, O.; KIRCHBERG, K. J.; FRISCHHOLZ, R. W. "Robust Face Detection Using the Hausdorff Distance" *Third International Conference on Audio- and Video-Based Biometric Person Authentication*, 2001: 90–95.

KOESTINGER, M.; WOHLHART, P.; ROTH, M.; BISCHOF, H. "Annotated Facial Landmarks in the Wild: A Large-scale, Real-world Database for Facial Landmark Localization" *IEEE International Workshop on Benchmarking Facial Image Analysis Technologies*, 2011.

KRUPPA, H.; CASTRILLÓN, M.; SCHIELE, B. "Fast and Robust Face Finding via Local Context" *IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance*, 2003: 157-164.

LANITIS, A.; TAYLOR, C.J.; COOTES, T.F. "Automatic Interpretation and Coding of Face Images Using Flexible Models" *IEEE Transaction Pattern Analysis and Machine Intelligence*, 1997: 743-756.

LIENHART, R.; KURANOV, A.; PISAREVSKY, V. "Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection" *Pattern Recognition Symposium*, 2003: 297-304.

LIENHART, R.; LIANG, L.; KURANOV, A. "A Detector Tree of Boosted Classifiers for Real-Time Object Detection and Tracking" *IEEE International Conference on Multimedia and Expo*, 2003: 277-280.

LIENHART, R.; MAYDT, J. "An extended set of Haar-like features for rapid object detection" *IEEE Computer Vision and Pattern Recognition*, 2002: 900-903.

LU, J.; PLATANIOTIS, K. N.; VENETSANOPOULOS, A.N. "Face Recognition Using LDA-Based Algorithms" *IEEE Transaction on Neural Networks*, 2003. Vol. 14, No. 1, 2003: 195-200.

MARSZALEC, E.; MARTINKAUPI, B.; SORIANO, M.; PIETKAINEN, M. "A Physics-based Face Database for Color Research". *Journal of Electronic Imaging*, 2000:32–38.

MARTINZES, R.; BENAVENTE, R. "The AR Face Database", *CVC Technical Report #24*. Junho 1998.

MESSER, K.; MATAS, J.; KITTLER, J. "Acquisition of a Large Database for Biometric Identity Verification" BIOSIGNAL 1998.

MOON, H; PHILLIPS, P. J. "Computational and Performance Aspects of PCA-based Face Recognition Algorithms", *Perception*, 2001. Vol. 30, 2001: 303-321.

OpenCV, Intel Open Source Computer Vision Library, <<http://sourceforge.net/projects/opencvlibrary>> Acesso em outubro 2011.

PADILLA, R.; COSTA FILHO, C. F. F.; COSTA, M. G. F. "Evaluation of Haar Cascade Classifiers for Face Detection" *International Conference on Digital Image Processing*, 2012. Vol. 1, 2012: 362-365.

PAPAGEORGIU, C.; OREN, M.; POGGIO, T. "A general framework for object detection" *International Conference on Computer Vision*, 1998.

PHILLIPS, P. J.; MOON, H.; RAUSS, P.J.; RIZVI, S. "The FERET Evaluation Methodology for Face Recognition Algorithms", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, Vol. 22, No. 10.

SAMARIA, F.; HARTER, A. "Parameterisation of a Stochastic Model for Human Face Identification" *2nd IEEE Workshop on Applications of Computer Vision*, 1994, Sarasota FL.

SCHMIDT, A.; KASINSKI, A. "The Performance of the Haar Cascade Classifiers Applied to the Face and Eyes Detection" *Advances in Soft Computing*, 2008, Vol. 45, 2007: 816-823.

SIM, T.; BAKER, S.; BSAT M. "The CMU Pose, Illumination, and Expression Database" *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2003, Vol. 25, No. 12, 2003: 1615-1618.

SOHAIL, A.S.M.; BHATTACHARYA P. "Detection of Facial Feature Points Using Anthropometric Face Model," *IEEE International Conference on Signal-Image Technology and Internet-Based Systems*, 2006: 656-665.

THOMAZ, C.; GIRALDI, G. "A new ranking method for principal components analysis and its application to face image analysis" *Journal Image and Vision Computing*, 2010. Vol. 28, No 6, 2010: 902-913.

TOOLE, A. O.; HARMS, S.; HURST, D. R.; PAPPAS, M.; ABDI, H."A Video Database of Moving Faces and People," 2003.

TURK, M.A.; PENTLAND, A.P. "Face Recognition Using Eigenfaces" *IEEE Conference on Computer Vision and Pattern Recognition*, 1991: 586-591.

VIOLA, P.; JONES, M. "Rapid Object Detection using a Boosted Cascade of Simple Features" *IEEE Computer Vision and Pattern Recognition*, 2001: 511-518.

VIOLA, P.; JONES, M. "Robust Real-time Object Detector", *International Journal of Computer Vision*, 2004: 137-154.

WILSON, P. I.; FERNANDEZ, J. "Facial feature detection using haar classifiers" *Journal of Computing Sciences in Colleges*, 2006: 127-133.

YUN, T.; GUAN, L. "Automatic Fiducial Points Detection for Facial Expressions Using Scale Invariant Feature" *IEEE International Workshop on Multimedia Signal Processing*, 2009: 1-6.

## APÊNDICE

Proceedings of ICDIP 2012: International Conference on Digital Image Processing, 2012. v. 1. p. 362-265.

# Evaluation of Haar Cascade Classifiers Designed for Face Detection

R. Padilla, C. F. F. Costa Filho and M. G. F. Costa

**Abstract**—In the past years a lot of effort has been made in the field of face detection. The human face contains important features that can be used by vision-based automated systems in order to identify and recognize individuals. Face location, the primary step of the vision-based automated systems, finds the face area in the input image. An accurate location of the face is still a challenging task. Viola-Jones framework has been widely used by researchers in order to detect the location of faces and objects in a given image. Face detection classifiers are shared by public communities, such as OpenCV. An evaluation of these classifiers will help researchers to choose the best classifier for their particular need. This work focuses of the evaluation of face detection classifiers minding facial landmarks.

**Keywords**—Face datasets, face detection, facial landmarking, haar wavelets, Viola-Jones detectors.

## I. INTRODUCTION

ALTHOUGH recognizing an individual by the face is an easy task for humans, it is a challenge for vision-based automated systems. It has been an active research area involving several disciplines such as image processing, neural networks, statistics, pattern recognition, anthropometry and computer vision. Vision-based automated systems can apply facial recognition and facial identification in numerous commercial applications, such as biometric authentication, human-computer interaction, surveillance, games and multimedia entertainment.

Unlike other biometrics, face recognition is non-invasive, and does not need physical contact of the individual with the system, making it a very acceptable biometric. Vision-based automated systems applied to face recognition can be divided into 4 steps: face detection, image pre-processing, feature extraction and matching [1]. Face detection is a hard task, once faces form a similar class of objects and their features, such as eyes, mouth, nose and chin, have, in general, the same geometrical configuration. The captured image of the face may be pre-processed to overcome illumination variations [2]. Feature extraction is the process where a geometrical or vectorial model is obtained gathering important characteristics

presented on the face. Feature extraction can be divided into 3 approaches: holistic, feature-based and hybrid. Principal component analysis [3] [4], fisher discriminant analysis [5] [6] and support vector machine [7] are examples of holistic approach. Feature-based approach is based on geometrical relation of the facial features. [8] applied active shape model, gathering important information presented in some of the facial features. Statistical classifiers such as Euclidian distance [9], Bayes classifier [10], Mahalanobis distance [11] and neural classifiers [12] can be used to compare the characteristic vector with other classes (individuals) in the matching step.

Face detection has been improved in terms of speed with the application of haar-features with the contribution of the Viola-Jones object detection framework. Implementations of this framework, such as OpenCV, provide different face classifiers created by authors that used different datasets into their training. The performance and reliability of these classifiers vary a lot. [13] evaluated the performance of some classifiers and also tested their accuracy.

This paper focuses on evaluating facial classifiers regarding facial features contained in the found face. We propose a method using different scores given to each facial feature contained in the located face. Two different face databases (FEI database and yale face database) were used to evaluate 10 face classifiers.

## II. MATERIALS

### A. Yale face database

The yale face database [14] contains facial images of 15 individuals, with 11 pictures per person, taken with different illumination conditions. The subjects have different facial expressions (with glasses, sad, sleepy, surprised, wink). The size of each image is 320x243 pixels.

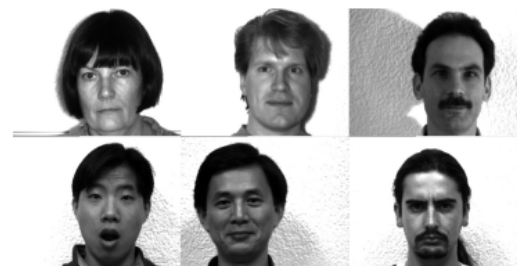


Fig. 1 Yale face database images

R. Padilla is with Universidade Federal do Amazonas and Nokia Institute of Technology, Av. Torquato Tapajós, 7200 – Col. Terra Nova, Manaus-AM Brazil. CEP: 69048-660 (phone: 55-92-2126-1000; e-mail: ext-rafael.padilla@nokia.com).

C. F. F. Costa Filho is with Universidade Federal do Amazonas, Av. Gal. Rodrigo Otavio Jordao Ramos, 3000, Manaus-AM Brazil (phone: 55-92-3305-4696; e-mail: cfffilho@gmail.com).

M. G. F. Costa is with Universidade Federal do Amazonas, Av. Gal. Rodrigo Otavio Jordao Ramos, 3000, Manaus-AM Brazil (phone: 55-92-3305-4696; e-mail: marly.costa@uol.com.br).

### B. FEI face database

The FEI face database [15] is a Brazilian database containing 14 images for each of 200 individuals, with a total of 2800 images. The images are colorful in different rotations with neutral, smiling and non-smiling expressions. We used 2 frontal images per individual, considering the smiling and non-smiling expression, in a total of 400 images. The original size of each image is 640x480 pixels.

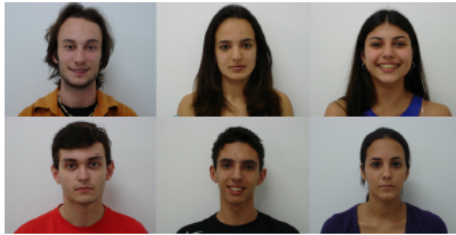


Fig. 2 FEI face database images

### C. Viola-Jones face detectors

Motivated by the challenge of face detection, [16] proposed an object detector framework using Haar-like features, which has been widely used by other works not only for face detection, but also for object locations.

Thanks to the Open Computer Vision Library implementation [17], the general object detector framework has become popular and motivated the community to generate their own object classifiers. These classifiers use haar-like features that are applied over the image. Only those image regions, called sub-windows, that pass through all the stages of the detector are considered to contain the target object. Fig. 3 shows the detection cascade schematic with  $N$  stages. The detection cascade is designed to eliminate a large number of negative examples with a little processing.

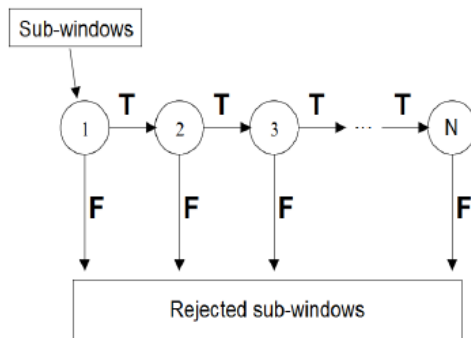


Fig. 3 Detection cascade

Some face location classifiers are distributed in the OpenCV implementation [17]. In this work, we used 10 classifiers that are presented in Table I. We kept the same description used by [13] to ease the comparisons.

TABLE I  
CLASSIFIERS

Classifier	Size	Stages	References	Target faces
FD	24x24	25	[18,19]	Frontal
FA1	20x20	22	[18, 19]	Frontal
FAT	20x20	46	[18, 19]	Frontal
FA2	20x20	20	[18, 19]	Frontal
FW	30x30	19	[20]	Frontal
FWQ	30x30	20	[20]	Quarter tuned
FWH	25x30	20	[20]	Half tuned
PR	20x20	26	[21]	Profile
HS1	22x18	30	[22]	Upper Body
HS2	22x20	19	[22]	Head and Shoulders

### III. LANDMARKS

Landmark detection is important not only to generate a geometric face model, but also can be used for face detection [23]. [24] compared different algorithms for facial landmark localization and proposed a set of tools that ease the integration of other face databases. [25] proposed a technique for face segmentation using Active Shape Model based on border landmarks of the face. [26] used a facial geometrical model based on the distance of the eyes to stipulate the position of other landmarks for face segmentation, shown in Fig. 4.

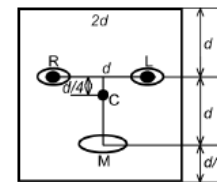


Fig. 4 Geometrical model of the face (Liu, Z et al – 2008)

FGnet project has published the location of 22 facial features of each face of the AR face database [27]. We also marked manually the same 22 facial feature points of the yale and FEI face database images used in this work. Fig. 5 shows an image with the marked facial points. In the total, 565 images were used and for each one of the 22 landmarks, a score was given (see Table II). The scores were either 1 or 2. The landmarks located in the contour of the face were given the highest score. The application of the scores will be explained in the next section.

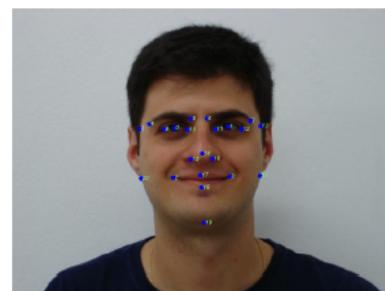


Fig. 5 Example of landmarks marked manually

TABLE II  
LANDMARKS AND SCORES

Landmark	Description	Score
0	Center of the right eye	1
1	Center of the left eye	1
2	Right corner of the mouth	1
3	Left corner of the mouth	1
4	Right eyebrow right corner	2
5	Right eyebrow left corner	2
6	Left eyebrow right corner	2
7	Left eyebrow left corner	2
8	Right upper facial limit	2
9	Right eye right corner	1
10	Right eye left corner	1
11	Left eye right corner	1
12	Left eye left corner	1
13	Left upper face limit	2
14	Nose tip	1
15	Right nostril	1
16	Left nostril	1
17	Upper lip limit	1
18	Bottom lip limit	1
19	Chin	2
20	Right facial limit	2
21	Left facial limit	2

#### IV. EXPERIMENTS

##### A. Preparation

First, we collected face location classifiers. They differ among each other by the number of stages and the minimal size of the faces that can be detected. They were designed to detect faces in different positions (frontal and profile faces) and accuracy (head only and head and shoulders together).

Two criteria were used to determine the precision of each classifier.

##### B. Criterium I

The accuracy of a classifier was measured by the score obtained by the located face. The detected face region must contain as many facial features as possible. The scores of each facial feature were added if the rectangle representing the position of the face overlaps these features.

Fig. 6 shows an example of a face with a face located by a classifier. Notice that there are 2 important feature points missing, landmarks 13 and 21.

By analyzing different faces, we estimated that a face is well located if its total score is higher or equal to 27, once lower scores leave important feature points out of the face image.

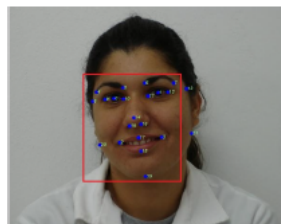


Fig. 6 Face located with a classifier, having score equals to 27

##### C. Criterium II

We also considered the same criterium used by [13] to evaluate the accuracy of the size of the detected face. A face is correctly detected if its height and width are not greater than four times the distance between the eyes. Differently from [13], we did not evaluate the time of processing of each classifier, our goal is to evaluate the accuracy of the classifiers regarding facial features and the size of the retrieved face.

#### V. RESULTS

By applying the selected classifiers in both databases and analyzing the precision of important facial features included in the detected face, described by Criterium I, it was noticed that the classifiers FA1, FAT and FA2 obtained better results. When FA1 and FA2 were used to locate faces in yale database, 100% of the faces obtained scores equal or higher than 27. FAT classifier obtained better results when used with FEI database, having 99.25% of the faces a score equal or higher than 27. Fig. 7 shows the percentage of images for both databases with score equal or higher than 27 of each classifier. Criterium II was used to determine if the region of the detected face is accurate. The best results with yale database were obtained by applying the classifiers FD, FA1 and FA2, where 100% of the images had faces detected with their height and width not greater than for times the distance of the eyes. With FEI database, the classifiers FAT, FWQ and FA1 had better results with percentages equal to 99.25%, 98.75% and 98.50% respectively.

It is important to mention that the images used in this work are frontal faces images and the classifiers designed to locate mainly frontal faces are FD, FA1, FAT, FA2 and FW. It explains the poor results obtained by the other classifiers.

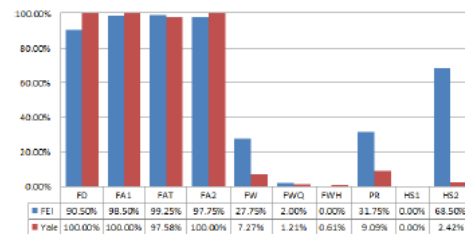


Fig. 7 Results using criterium I

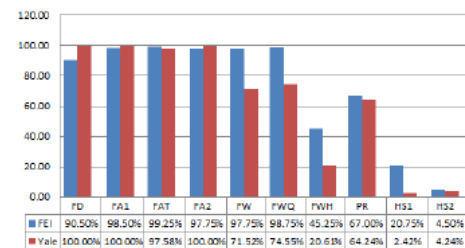


Fig. 8 Results using criterium II



## REFERENCES

- [1] J. Fagertun, 2005. Face Recognition. Master Thesis, Technical University of Denmark (DTU).
- [2] Y. Gang, L. Jiawei, L. Jiayu, M. Qingli and Y. Ming, "Illumination Variation in Face Recognition: A Review", *IEEE Second International Conference on Intelligent Networks and Intelligent Systems (ICINIS 2009)*, pp. 309-311.
- [3] M.A. Turk, A.P. Pentland, "Face Recognition Using Eigenfaces", Proceedings of the *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'91)*, 3-6 June 1991, Maui, Hawaii, USA, pp. 586-591.
- [4] H. Moon, P.J. Phillips, "Computational and Performance Aspects of PCA-based Face Recognition Algorithms", *Perception*, Vol. 30, 2001, pp. 303-321.
- [5] K. Etemad, R. Chellappa, "Discriminant Analysis for Recognition of Human Face Images", *Journal of the Optical Society of America A*, Vol. 14, No. 8, August 1997, pp. 1724-1733.
- [6] J. Lu, K.N. Plataniotis, A.N. Venetsanopoulos, "Face Recognition Using LDA-Based Algorithms", *IEEE Transaction on Neural Networks*, Vol. 14, No. 1, January 2003, pp. 195-200.
- [7] B. Heisele, P. Ho, T. Poggio, "Face Recognition with Support Vector Machines: Global versus Component-based Approach", Proceedings of the *Eighth IEEE International Conference on Computer Vision (ICCV'01)*, Vol. 2, 09-12 July 2001, Vancouver, Canada, pp. 688-694.
- [8] A. Lanitis, C.J. Taylor, T.F. Cootes, "Automatic Interpretation and Coding of Face Images Using Flexible Models", *IEEE Transaction Pattern Analysis and Machine Intelligence* (1997), pp. 743-756.
- [9] E. Gomathi, K. Baskaran, "Recognition of Faces Using Improved Principal Component Analysis", *Second International Conference on Machine Learning and Computing (ICMLC'10)*, pp.198-201.
- [10] C. Liu and H. Wechsler, "Probabilistic Reasoning Models for Face Recognition", in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'98)*, pp.827-832.
- [11] Y. Ji, T. Lin, and H. Zha, "Mahalanobis Distance Based Non-negative Sparse Representation for Face Recognition", in Proceedings *The Eighth International Conference on Machine Learning and Applications (ICMLA'09)*, pp.41-46.
- [12] V. Kabeer & N. K. Narayanan, "Face recognition using state space parameter and Artificial Neural Network Classifier", Proceedings of *IEEE International Conference on Computational Intelligence and Multimedia Applications (ICCI'07)*, Sivakasi, India Vol.3, December, 2007, pp 250-254.
- [13] M. Castrillón, O. Deniz, D. Hernández, and J. Lorenzo, "A Comparison of Face and Facial Feature Detectors based on the Viola-Jones General Object Detection Framework", *Machine Vision and Applications*, vol. 22 issue 3, 2011.
- [14] P. N. Bellhumer, J. Hespanha, and D. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection", *IEEE Transactions on Pattern Analysis and Machine Intelligence, Special Issue on Face Recognition*, 1997, pp. 711-720.
- [15] C. Thomaz and G. Giraldi, "A new ranking method for principal components analysis and its application to face image analysis", *Journal Image and Vision Computing*, 2010, vol. 28, no. 6, pp. 902-913.
- [16] P. Viola and M. Jones, "Robust real-time object detection," *International Journal of Computer Vision*, 2002 vol. 57, no. 2, pp. 137-154.
- [17] Inte, Intel Open Source Computer Vision Library, v1.10re, <http://sourceforge.net/projects/opencvlibrary/> (October 2011).
- [18] Lienhart, R., Kuranov, A., Pisarevsky, V., "Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection", in *25th Pattern Recognition Symposium (DAGM'03)*, pp. 297-304.
- [19] Lienhart, R., Liang, L., Kuranov, A., "A Detector Tree of Boosted Classifiers for Real-Time Object Detection and Tracking", in *IEEE International Conference on Multimedia and Expo (ICME'03)*, pp. 277-280.
- [20] Wilson, P. I., Fernandez, J., "Facial feature detection using haar classifiers", in *Journal of Computing Sciences in Colleges* (2006), pp. 127-133.
- [21] Bradley, D. "Profile face detection" <http://www.cs.cmu.edu/~dbradley/publications/bradley-irac-03.swf>, (2003) last accessed 12/10/2011.
- [22] Kruppa, H., Castrillón Santana, M., Schiele, B., "Fast and Robust Face Finding via Local Context", in *IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS'03)*, pp. 157-164.
- [23] G. M. Beumer, Q. Tao, A. M. Bazen and R. N. J. Veldhuis, "A Landmark Paper in Face Recognition", *IEEE International Conference on Automatic Face and Gesture Recognition (FGE'02)*, pp. 73-78.
- [24] M. Koestinger, P. Wohlhart, P. M. Roth and H. Bischof, "Annotated Facial Landmarks in the Wild: A Large-scale, Real-world Database for Facial Landmark Localization", *IEEE International Workshop on Benchmarking Facial Image Analysis Technologies (BeFIT'11)*.
- [25] M. Jian-Wei and F. Yu-Hua, "Face segmentation algorithm based on ASM", *IEEE Conference on Intelligent Computing and Intelligent Systems (ICIS'09)*, 2009, pp 495-499.
- [26] Z. Liu, W. Li, X. Zhang and J. Yang, "Efficient Face Segmentation Based on Face Attention Model and Seeded Region Merging", *9th International Conference on Signal Processing (ICSP'08)*, pp. 1116-1119.
- [27] A.M. Martinez and R. Benavente, "The AR Face Database", CVC Technical Report #24, June - 1998.

# Avaliação de Classificadores Haar Projetados para Detecção Facial

R. Padilla<sup>1</sup>, C. F. F. Costa Filho<sup>2</sup>, M. G. F. Costa<sup>2</sup>

<sup>1</sup>Instituto Nokia de Tecnologia

Av. Torquato Tapajós, 7200 – Col. Terra Nova, Manaus-AM. CEP: 69048-660 Brasil

<sup>2</sup>Universidade Federal do Amazonas

Av. Gal. Rodrigo Otávio Jordão Ramos, 3000, Manaus-AM. CEP: 69010-110 Brasil

rafael.padilla@indt.org.br, cffcfilho@gmail.com, marly.costa@uol.com.br

**RESUMO** – Nos últimos anos vários estudos foram realizados na área de detecção de faces. A face humana contém características importantes que podem ser utilizadas em sistemas biométricos para a identificação de indivíduos. A localização facial, a primeira etapa de sistemas biométricos pela face, localiza faces na imagem de entrada. O *framework* Viola-Jones tem sido amplamente utilizado por pesquisadores para detecção e localização de objetos e faces. Classificadores de detecção facial são compartilhados pela comunidade científica e acadêmica. Este trabalho tem como objetivo avaliar classificadores localizadores de face através de pontos faciais, auxiliando pesquisadores na escolha do melhor classificador de acordo com a necessidade de cada um.

## 1. INTRODUÇÃO

Embora o reconhecimento de um indivíduo através da face seja uma tarefa trivial para seres humanos, é um desafio para sistemas de automação baseados em imagens. Este tema tem sido estudado por pesquisadores de diferentes áreas envolvendo processamento digital de imagens, redes neurais, estatística, reconhecimento de padrões, antropometria e visão computacional. Sistemas de automação baseados em imagens fazem o reconhecimento e identificação facial em diversas aplicações comerciais, tais como autenticação por biometria, interação homem-máquina, vigilância, jogos e entretenimento.

Ao contrário de outras biometrias, o reconhecimento de face é não-invasivo e não necessita de contato físico do indivíduo com o sistema, tornando-se uma tecnologia muito aceitável no mercado. Sistemas automáticos baseados em visão para o reconhecimento de indivíduos pela face pode ser dividido em 4 etapas: detecção da face, pré-processamento da imagem, extração de características e reconhecimento [1].

A imagem capturada da face pode ser pré-processada para eliminar as variações de iluminação [2]. A extração de características é o processo onde um modelo geométrico ou vetorial é obtido reunindo características importantes presentes na face. A extração de características pode ser dividida em 3 abordagens: holística, baseada em características e híbrida. São exemplos de abordagens holísticas: Análise de Componentes Principais [3] [4], Análise Discriminante de Fisher [5] [6] e Máquina de Vetores de Suporte [7]. Extrair características de uma determinada face é uma tarefa difícil, uma vez que suas características, tais como olhos, boca, nariz e queixo, possuem, em geral, a mesma configuração geométrica. Abordagens baseadas em características utilizam as relações geométricas das características faciais. O Modelo de Forma Ativa pode ser aplicado para obter informações importantes presentes em alguns pontos faciais [8]. Classificadores estatísticos, tais como Distância Euclidiana [9], Classificador de Bayes [10], Distância de Mahalanobis [11] e Classificadores Neurais [12] podem ser usados na comparação do vetor de características com outras classes (indivíduos) durante a etapa de reconhecimento.

A detecção de faces teve sua acurácia melhorada com a aplicação de características tipo Haar, com a contribuição do *framework* de detecção de objetos Viola-Jones. Implementações deste *framework* tornaram-se populares e proporcionam a criação de classificadores por autores que utilizaram diferentes conjuntos de imagens para sua

geração. O desempenho e a confiabilidade desses classificadores varia muito [13].

Este artigo tem como objetivo avaliar classificadores faciais considerando características faciais contidas nas faces encontradas. Propomos um método utilizando diferentes pontuações dadas para cada característica facial contida na face localizada. Dois bancos de dados de faces (FEI e Yale) foram utilizados para avaliar 10 classificadores faciais.

## 2. MATERIAIS

### 2.1. BASE DE FACES YALE

A base de faces Yale [14] contém imagens dos rostos de 15 indivíduos, com 11 fotos cada um, tiradas em diferentes condições de iluminação. Os indivíduos apresentam expressões faciais variadas (com óculos, triste, sonolento, piscando, surpresos). O tamanho de cada imagem é 320x243 pixels.

A Figura 1 mostra exemplos de imagens da base de faces Yale.

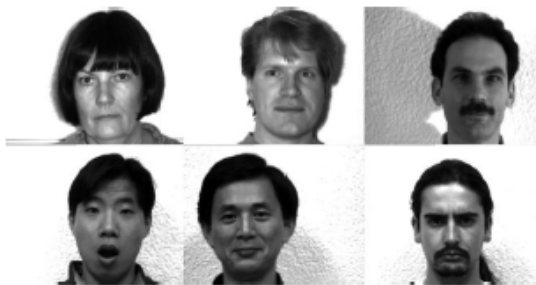


Figura 1 – Base de faces Yale

### 2.2. BASE DE FACES FEI

A base de faces FEI [15] é um banco de dados brasileiro contendo 14 imagens para cada um dos 200 indivíduos, com um total de 2800 imagens. As imagens são coloridas e apresentam faces em diferentes rotações, com expressões neutras e sorrindo. Utilizamos 2 imagens frontais por indivíduo, considerando as imagens com sorriso e neutras, totalizando 400 imagens. O tamanho original de cada imagem é 640x480 pixels. Exemplos de imagens da base de faces FEI podem ser vistos na Figura 2.



Figura 2 – Base de faces FEI

### 2.3. DETECTORES DE FACE VIOLA-JONES

Motivados pelo desafio de detecção de faces, [16] propuseram um *framework* detector objetos utilizando características tipo Haar. Este *framework* tem sido amplamente utilizado por outros pesquisadores e trabalhos, não só para a detecção de faces, mas também para localização de objetos em imagens e vídeos.

Graças à implementação da biblioteca OpenCV [17], o *framework* Viola-Jones tornou-se popular e motivou a comunidade a gerar seus próprio classificadores de objetos. Esses classificadores usam características tipo Haar que são aplicadas sobre a imagem.

O conceito de características tipo Haar provém de funções Haar de Base [18]. Existem basicamente 4 tipos de características, que são demonstradas na Figura 3.

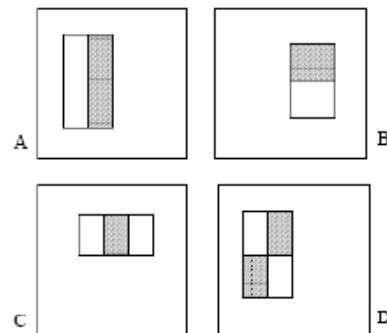


Figura 3 – Características tipo Haar posicionadas sobre uma janela de amostragem. FONTE: P. Viola and M. Jones [16]

Cada característica posicionada sobre uma parte da imagem, chamada de sub-janela, expressa um valor obtido através da soma dos pixels da região branca subtraídos da soma dos pixels da região escura.

Uma vez obtido o conjunto de características e seus limiares, o processo de teste é feito utilizando uma técnica de verificação em cascata, conforme a Figura 4. Nesta figura temos um classificador formado por  $N$  camadas. Cada camada é composta por várias características. Para o classificador comprovar que uma sub-janela possua uma face, esta sub-janela deve passar por todas as  $N$  camadas do classificador, de modo que não seja rejeitada por nenhuma delas.

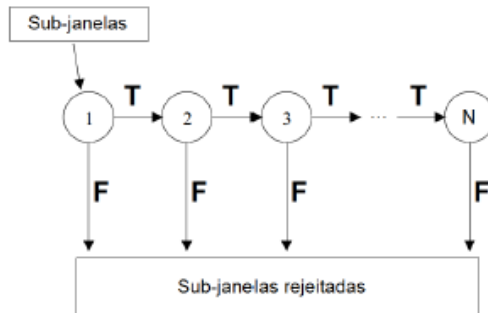


Figura 4 – Detecção em cascata

Alguns classificadores de localização de faces são distribuídos na implementação do OpenCV [17]. Estes classificadores são arquivos XML (Extensible Markup Language) estruturados em estágios como mostra a Figura 5.



Figura 5 – Estrutura de um classificador Viola-Jones em XML

Cada classificador é formado por vários estágios. O classificador determina o tamanho mínimo da sub-janela que será analisada pelos estágios até que seja reprovada por um deles. Caso a sub-janela seja reprovada, a mesma é deslocada pela imagem, gerando uma nova sub-janela a ser analisada pelo classificador. Ao passar por toda imagem, a sub-janela terá suas dimensões ampliadas e o processo

recomeça através de toda a imagem, conforme mostra a Figura 6.



Figura 6 – Sub-janela deslocada através de uma imagem

Cada estágio é formado por características e por um limiar de estágio. Cada característica é formada por retângulos, por um limiar de característica, valor esquerdo e valor direito. A Figura 7 mostra a estrutura de um único estágio.

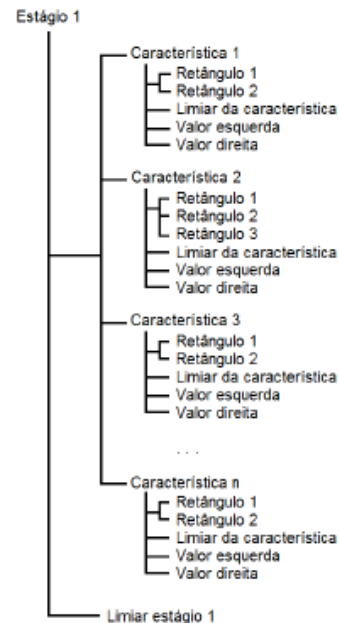


Figura 7 – Estrutura de um estágio utilizado no classificador de objetos

Cada característica possui um valor esquerda e um valor direita. Estes valores são pesos encontrados no processo de treinamento e atribuídos a cada característica, que são somados e comparados com o limiar do estágio.

O valor de cada característica é calculado pela aplicação da mesma na sub-janela. Se o valor da característica for menor do que seu limiar, acumula-se o valor da esquerda, caso contrário, o valor acumulado é o valor da direita. Para uma sub-janela ser aprovada por um estágio, o valor acumulado por todas as características deve ser menor do que o limiar do estágio.

Neste trabalho, foram utilizados 10 classificadores, apresentados na Tabela I. Mantivemos a mesma descrição utilizada por [13] para facilitar comparações.

TABELA I  
CLASSIFICADORES

Classificador	Tamanho	Estágios	Referências	Posição da face
FD	24x24	25	[19,20]	Frontal
FA1	20x20	22	[19,20]	Frontal
FAT	20x20	46	[19,20]	Frontal
FA2	20x20	20	[19,20]	Frontal
FW	30x30	19	[21]	Frontal
FWQ	30x30	20	[21]	Rotacionada $\frac{1}{4}$
FWH	25x30	20	[21]	Rotacionada $\frac{1}{2}$
PR	20x20	26	[22]	Perfil
HS1	22x18	30	[23]	Parte superior do corpo
HS2	22x20	19	[24]	Cabeça e ombros

Selecionamos o maior número de classificadores disponíveis pela comunidade científica. A escolha desses classificadores resultou de uma pesquisa detalhada realizada na literatura, de onde extraímos as referências listadas na Tabela 1 e seus correspondentes classificadores. Os classificadores FW, FWQ, FWH, HS1 e HS2 foram gentilmente cedidos pelos autores através de contato realizado. Os classificadores PR, FD, FAT, FA2, FA1 estão disponíveis na versão 2.4 da biblioteca OpenCV.

### 3. PONTOS FACIAIS

A detecção de pontos faciais é importante não somente para gerar um modelo geométrico do rosto, mas também para ser utilizado na detecção de faces [24]. [25] compararam diferentes algoritmos para a localização de

pontos faciais e propuseram um conjunto de ferramentas que facilitam a integração com outras bases de dados. [26] propuseram uma técnica para segmentação facial usando Modelo de Forma Ativa com base em pontos de fronteira do rosto. [27] utilizou um modelo geométrico facial com base na distância dos olhos para estipular a posição dos pontos da face. Este modelo geométrico é mostrado na Figura 8.

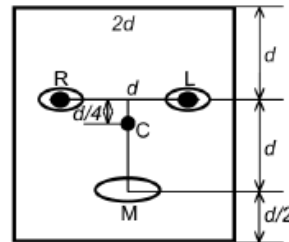


Figura 8 – Detecção em cascata

O projeto FGnet publicou a localização de 22 traços faciais de cada face do banco de faces AR [28]. Nós também marcamos manualmente estes 22 pontos faciais nas bases de Yale e FEI utilizadas neste trabalho. A Figura 9 mostra uma imagem com os pontos faciais marcados.

No total, 565 imagens foram utilizados e para cada um dos 22 pontos de referência, uma pontuação foi dada (ver Tabela 2). A pontuação dada para cada ponto foi 1 ou 2. Para pontos localizados no contorno da face a maior pontuação foi dada. A aplicação das pontuações será explicado na sessão seguinte.

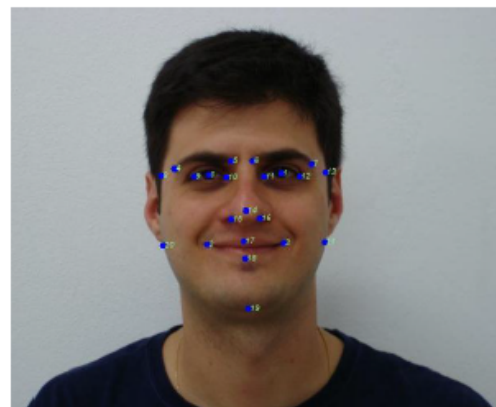


Figura 9 – Exemplo de pontos faciais marcados manualmente

TABELA 2  
PONTOS FACIAIS E SUAS PONTUAÇÕES

Ponto facial	Descrição	Pontuação
0	Centro do olho direito	1
1	Centro do olho esquerdo	1
2	Canto direito da boca	1
3	Canto esquerdo da boca	1
4	Canto direito da sobrancelha direita	2
5	Canto esquerdo da sobrancelha direita	2
6	Canto direito da sobrancelha esquerda	2
7	Canto esquerdo da sobrancelha esquerda	2
8	Límite superior direito da face	2
9	Canto direito do olho direito	1
10	Canto esquerdo do olho direito	1
11	Canto direito do olho esquerdo	1
12	Canto esquerdo do olho esquerdo	1
13	Límite superior esquerdo da face	2
14	Ponta do nariz	1
15	Narina direita	1
16	Narina esquerda	1
17	Límite superior labial	1
18	Límite inferior labial	1
19	Queixo	2
20	Límite direito da face	2
21	Límite esquerdo da face	2

#### 4. EXPERIMENTOS

Primeiramente coletamos 10 classificadores de localização facial. Eles diferenciam-se pelo número de estágios e pelo tamanho mínimo de faces que podem ser detectadas. Eles foram projetados para detectar rostos em diferentes posições (frontal e de perfil) e somente a cabeça ou cabeça e ombros e juntos.

Dois critérios foram utilizados para determinar a precisão de cada classificador, e são descritos a seguir.

##### 4.1. CRITÉRIO I

A precisão de um classificador foi mensurada pela pontuação obtida pelas faces localizadas. A região do rosto detectado deve conter o máximo de pontos faciais possível. A pontuação de cada característica facial é somada se esta encontra-se dentro do retângulo que representa a face.

A Figura 10 mostra o exemplo de uma face localizada por um dado classificador. Note que existem 2 pontos faciais importantes que não

estão dentro da face localizada - pontos 13 e 21.

Ao analisarmos diferentes faces, estimamos que um rosto é bem localizado, se sua pontuação total for maior ou igual a 27. Faces com pontuações mais baixas deixam importantes pontos faciais fora da imagem do rosto.

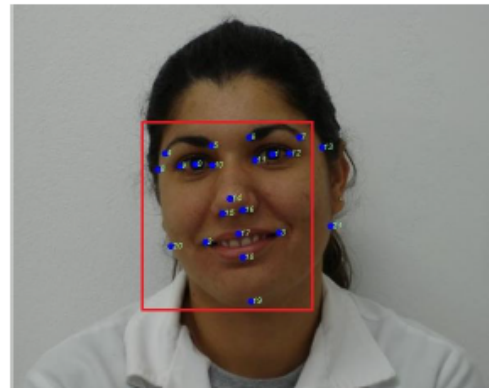


Figura 10 – Face localizada com pontuação igual à 27

##### 4.2. CRITÉRIO II

Nós também consideramos o mesmo critério utilizado por [13] para avaliar a precisão do tamanho da face detectada. Uma face é corretamente detectada se sua altura e largura são menores que quatro vezes a distância entre os olhos. Diferentemente de [13], não avaliamos o tempo de processamento de cada classificador. Nosso objetivo é avaliar a acurácia dos classificadores, considerando pontos faciais (critério I) e o tamanho das faces localizadas (critério II).

##### 4.3. RESULTADOS

Podemos ver os resultados da aplicação dos dez classificadores nas duas bases nas Figuras 11 e 12.

A Figura 11 mostra os resultados obtidos quando o critério I foi utilizado e a Figura 12 com a utilização do critério II.

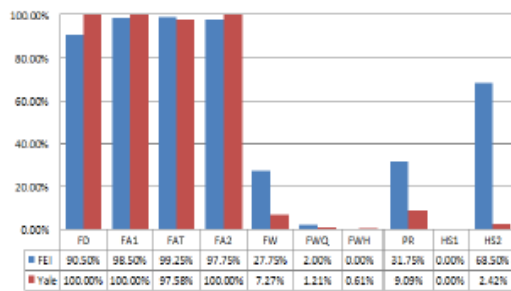


Figura 11 – Resultados utilizando o critério I

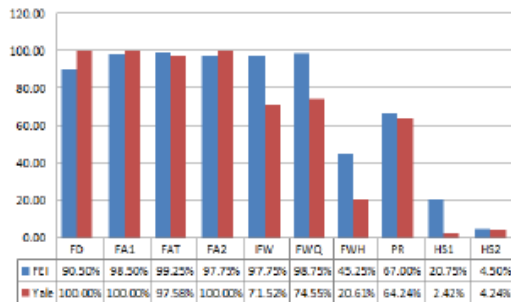


Figura 12 – Resultados utilizando o critério II

A partir dos dados obtidos, podemos concluir que:

a) Os melhores índices de aprovação foram alcançados pelos classificadores FA1, FA2 e FAT;

b) Nota-se que o índice de 100% de aprovação só foi conseguido em alguns classificadores, utilizando-se a base de imagens Yale;

c) Os classificadores FD, FA1, FAT, FA2 tiveram índices de aprovação semelhantes ao se utilizar os Critérios I e II, em ambas as bases;

d) Os classificadores FW, FWQ, FWH, PR e HS1 tiveram índices de aprovação bem menores, em ambas as bases, ao se utilizar o Critério I do que o Critério II;

e) O classificador HS2 teve um índice de aprovação maior na Base FEI ao se utilizar o Critério I, enquanto que o mesmo teve um índice de aprovação maior na Base Yale ao se utilizar o Critério II.

Assim sendo, a partir das conclusões anteriores podemos afirmar que os melhores classificadores avaliados foram FA1, FA2 e FAT. Em termos dos percentuais de

aprovação, o critério I, proposto neste trabalho, foi, na maioria dos casos, mais exigente que o critério II. A única exceção a essa regra foi verificada com o classificador HS2, quando avaliado na base FEI.

## REFERÊNCIAS

- [1] J. Fagertun, 2005. Face Recognition. Dissertação de Mestrado, Universidade Técnica da Dinamarca (DTU).
- [2] Y. Gang, L. Jiswei, L. Jiayu, M. Qingli and Y. Ming, "Illumination Variation in Face Recognition: A Review", *IEEE Second International Conference on Intelligent Networks and Intelligent Systems (ICINIS 2009)*, pp. 309-311.
- [3] M.A. Turk, A.P. Pentland, "Face Recognition Using Eigenfaces", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'91)*, 3-6 Junho 1991, Maui, Hawaii, USA, pp. 586-591.
- [4] H. Moon, P.J. Phillips, "Computational and Performance Aspects of PCA-based Face Recognition Algorithms", *Perception*, Vol. 30, 2001, pp. 303-321.
- [5] K. Etemad, R. Chellappa, "Discriminant Analysis for Recognition of Human Face Images", *Journal of the Optical Society of America A*, Vol. 14, No. 8, Agosto 1997, pp. 1724-1733.
- [6] J. Lu, K.N. Plataniotis, A.N. Venetsanopoulos, "Face Recognition Using LDA-Based Algorithms", *IEEE Transaction on Neural Networks*, Vol. 14, No. 1, Janeiro 2003, pp. 195-200.
- [7] B. Heisele, P. Ho, T. Poggio, "Face Recognition with Support Vector Machines: Global versus Component-based Approach", *Proceedings of the Eighth IEEE International Conference on Computer Vision (ICCV'01)*, Vol. 2, 09-12 Julho 2001, Vancouver, Canadá, pp. 688-694.
- [8] A. Lanitis, C.J. Taylor, T.F. Cootes, "Automatic Interpretation and Coding of Face Images Using Flexible Models", *IEEE Transaction Pattern Analysis and Machine Intelligence* (1997), pp. 743-756.
- [9] E. Gomathi, K. Baskaran, "Recognition of Faces Using Improved Principal Component Analysis", (ICMLC'10), pp.198-201.
- [10] C. Liu and H. Wechsler, "Probabilistic Reasoning Models for Face Recognition", in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'98)*, pp.827-832.
- [11] Y. Ji, T. Lin, and H. Zha, "Mahalanobis Distance Based Non-negative Sparse Representation for Face Recognition", em *The Eighth International Conference on Machine Learning and Applications (ICMLA'09)*, pp.41-46.
- [12] V. Kabeer & N. K. Narayanan, "Face recognition using state space parameter and Artificial Neural Network Classifier", *IEEE International Conference on Computational Intelligence and Multimedia Applications (ICCI'07)*, Sivakasi, Índia Vol.3, Dezembro, 2007, pp 250-254.
- [13] M. Castrillón, O. Déniz, D. Hernández, and J. Lorenzo. "A Comparison of Face and Facial Feature Detectors based on the Viola-Jones General Object Detection Framework". *Machine Vision and Applications*, vol. 22 issue 3, 2011.
- [14] P. N. Bellhumer, J. Hespanha, and D. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection", *IEEE Transactions on Pattern Analysis and Machine Intelligence, Special Issue on Face Recognition*, 1997, pp. 711-720.
- [15] L. L. de Oliveira Junior and C. E. Thomaz. "Captura e Alinhamento de Imagens: Um Banco de Faces Brasileiro". *Relatório Final de Graduação*. Departamento de Engenharia Elétrica, FEI, São Bernardo do Campo, São Paulo, Brasil, Junho 2006.
- [16] P. Viola and M. Jones, "Robust real-time object detection," *International Journal of Computer Vision*, 2002 vol. 57, no. 2, pp. 137-154.
- [17] Intel, Intel Open Source Computer Vision Library, v1.10re, <http://sourceforge.net/projects/opencvlibrary> (Outubro 2011).
- [18] PAPAGEORGIOU, C.; OREN, M.; POGGIO, T. "A general framework for object detection" *International Conference on Computer Vision*, 1998.

- [19] Lienhart, R., Kuranov, A., Pisarevsky, V., "Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection", em *25th Pattern Recognition Symposium (DAGM'03)*, pp. 297-304.
- [20] Lienhart, R., Liang, L., Kuranov, A., "A Detector Tree of Boosted Classifiers for Real-Time Object Detection and Tracking", em *IEEE International Conference on Multimedia and Expo (ICME'03)*, pp. 277-280.
- [21] Wilson, P. I., Fernandez, J., "Facial feature detection using haar classifiers", em *Journal of Computing Sciences in Colleges* (2006), pp. 127-133.
- [22] Bradley, D. "Profile face detection" <http://www.cs.cmu.edu/~dbradley/publications/bradley-irac-03.swf>, (2003) ultimo acesso em 10/12/2011.
- [23] Kruppa, H., Castrillón Santana, M., Schiele, B., "Fast and Robust Face Finding via Local Context", em *IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS'03)*, pp. 157-164.
- [24] G. M. Beumer, Q. Tao, A. M. Bazen and R. N. J. Veldhuis, "A Landmark Paper in Face Recognition". *IEEE International Conference on Automatic Face and Gesture Recognition (FG'02)*, pp. 73-78.
- [25] M. Koestinger, P. Wohlhart, P. M. Roth and H. Bischof, "Annotated Facial Landmarks in the Wild: A Large-scale, Real-world Database for Facial Landmark Localization", *IEEE International Workshop on Benchmarking Facial Image Analysis Technologies (BeFIT'11)*.
- [26] M. Jian-Wei and F. Yu-Hua, "Face segmentation algorithm based on ASM", *IEEE Conference on Intelligent Computing and Intelligent Systems (ICIS'09)*, 2009, pp 495-499.
- [27] Z. Liu, W. Li, X. Zhang and J. Yang, "Efficient Face Segmentation Based on Face Attention Model and Seeded Region Merging", *9th International Conference on Signal Processing (ICSP'08)*. pp. 1116-1119.
- [28] A.M. Martinez and R. Benavente, "The AR Face Database", CVC Technical Report #24, Junho - 1998.