

VIOLA JONES

Face Detection Algorithm

PROCESSAMENTO DIGITAL DE IMAGENS

MARCELO ARDIZZON LOVATTE



INTRODUÇÃO

- ROBUST REAL-TIME FACE DETECTION (2004)
 - Paul Viola;
 - Michael J. Jones;
- Proposta do artigo: *"a face detection framework that is capable of processing images extremely rapidly while achieving high detection rates"*;
- 384 by 288 pixel images;
- 15 frames per second on a conventional 700 MHz Intel Pentium III.

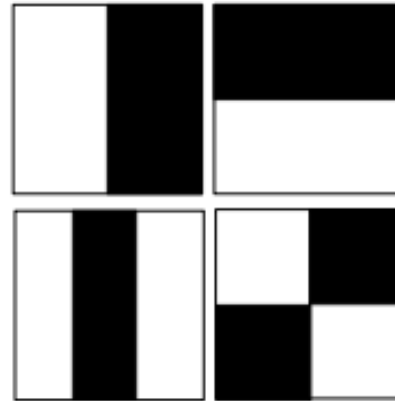
ALGORITIMO

1. Haar features;
2. Integral Image;
3. Adaboost;
4. Cascade.

HAAR FEATURES

Features retangulares

- Dois retângulos
- Três retângulos;
- Quatro retângulos.



Cada *feature* resulta em um valor que é calculado através da diferença entre o somatório da região preta com o somatório da região branca.

$$f(w) = \sum^w p_{preto} - \sum^w p_{branco}.$$

HAAR FEATURES

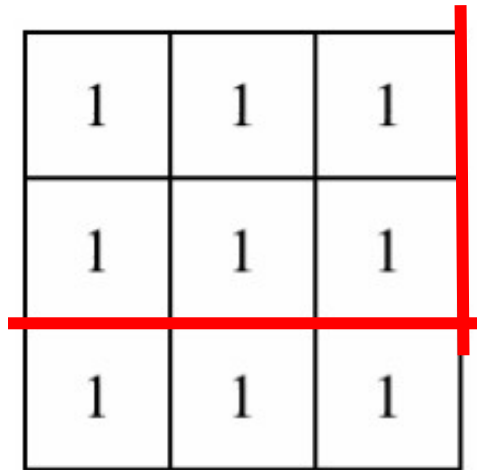
Haar features são utilizados para detectar a presença da mesma característica na imagem de entrada;



Considerando todos os tipos e parâmetros possíveis do *Haar features*, existem mais de **160.000** combinações.

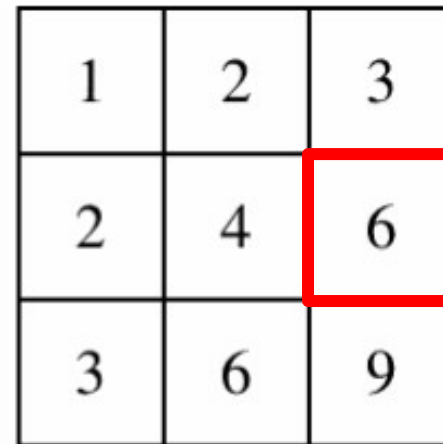
INTEGRAL IMAGE

Em uma integral imagem o valor do pixel (x,y) será o valor da soma dos pixels acima e à esquerda de (x,y) ;



1	1	1
1	1	1
1	1	1

Input image

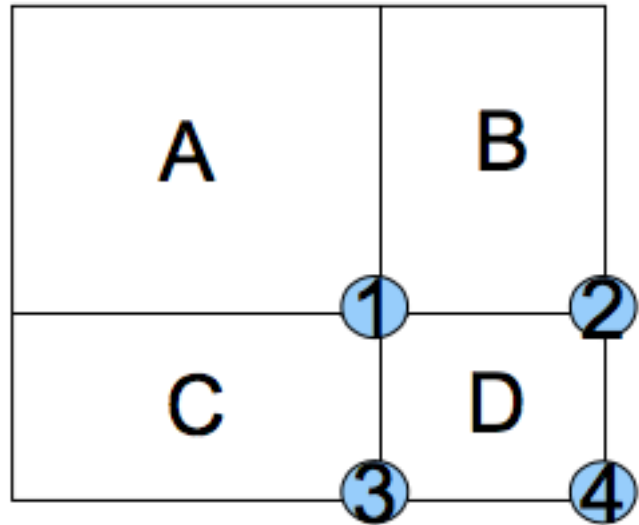


1	2	3
2	4	6
3	6	9

Integral image

INTEGRAL IMAGE

Utilizando a integral imagem, qualquer retângulo pode ser somado utilizando apenas quatro valores de referência na diagonal:



$$D = (1+4) - (2+3)$$

$$D = [A + (A+B+C+D)] - [A+B+A+C]$$

$$D = 2A+B+C+D - 2A - B - C$$

$$D = D$$

ADABOOST

160.000 *features* retangulares associados a cada sub-janela;

Apesar do conceito de *integral image* acelerar o processamento de cada *feature*, ainda se tem um grande volume de dados;

Hipótese: "*is that a very small number of these features can be combined to form an effective classifier*";

Utiliza-se uma variação do Adaboost para selecionar os *features* e treinar o classificador (Freund and Schapire, 1995);

O objetivo do algoritmo é selecionar os *features* retangulares que melhor separam os exemplos positivos dos negativos (faces e não-faces).

ADABOOST

Para cada *feature*, o algoritmo determina a função de classificação com o *threshold* ideal, baseado em um número mínimo de erros;

A função do classificador consiste em $h(x, f, p, \theta)$:

- x = sub-janela (24x24);
- f = *feature*;
- θ = *threshold*;
- p = polaridade;

Na prática, cada *feature* isolado não executa a classificação com baixo erro (10-50%).

Adaboost constrói um classificador forte através de uma combinação linear de fracos classificadores;

ADABOOST

200 *features* seleccionados;

Taxa de detecção de 95% - taxa falso positivo de 1 em 14084 no conjunto de dados de teste;

Apesar de promissor, um detector de face para aplicações reais, deve ter uma taxa de falsos positivos próximo de 1 em 1.000.000;

Em termos computacionais, o classificador é bem rápido, 0.7 segundos para escanear uma imagem 384 x 288;

Para melhorar a taxa de detecção, deve-se adicionar mais *features* ao classificador, aumentando diretamente o tempo de processamento.

CASCADE

Para aumentar a performance de detecção e reduzir o tempo computacional, utiliza-se um algoritmo de classificadores em cascata;

Classificadores mais simples são utilizados para rejeitar a maioria das sub-janelas antes que os classificadores mais complexos sejam utilizados para obter baixas taxas de falsos positivos;

Descartar não-faces rapidamente e "gastar" mais tempo em prováveis faces;

Os estágios do *cascade* são construídas através do classificadores de treinamento do Adaboost.

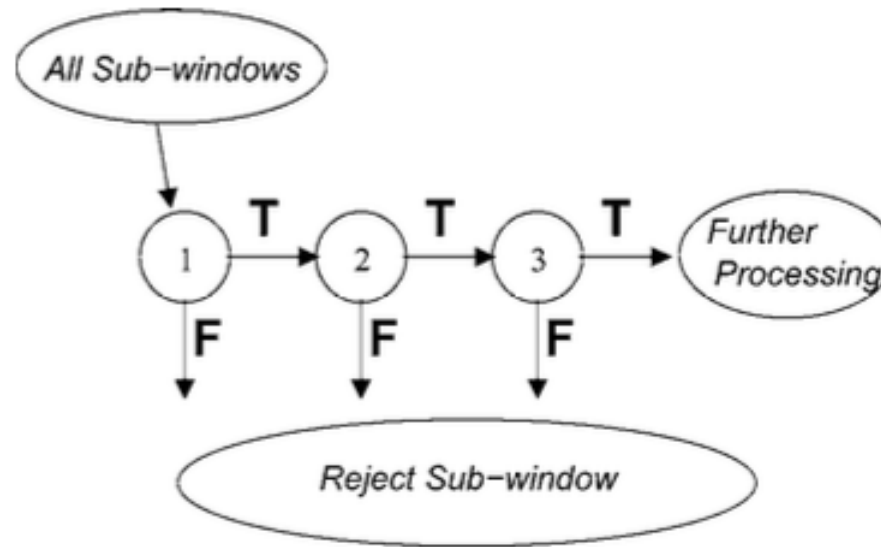
CASCADE

Primeiro estágio:

- Classificador forte com 2 *features*;
- *Threshold* ajustado para minimizar falsos negativos (descartar uma face);
- Detecção de 100% das faces com 50% de falsos positivos:

Estágios subsequentes:

- Eliminar as não-faces;



CASCADE

O detector final utilizou 38 camadas de classificadores e um total de 6060 *features*;

O **primeiro classificador** utilizou 2 *features*, rejeitou 50% das não-faces e detectou aproximadamente 100% das faces;

O **segundo classificador** utilizou 10 *features*, rejeitou 80% das não-faces e detectou aproximadamente 100% das faces;

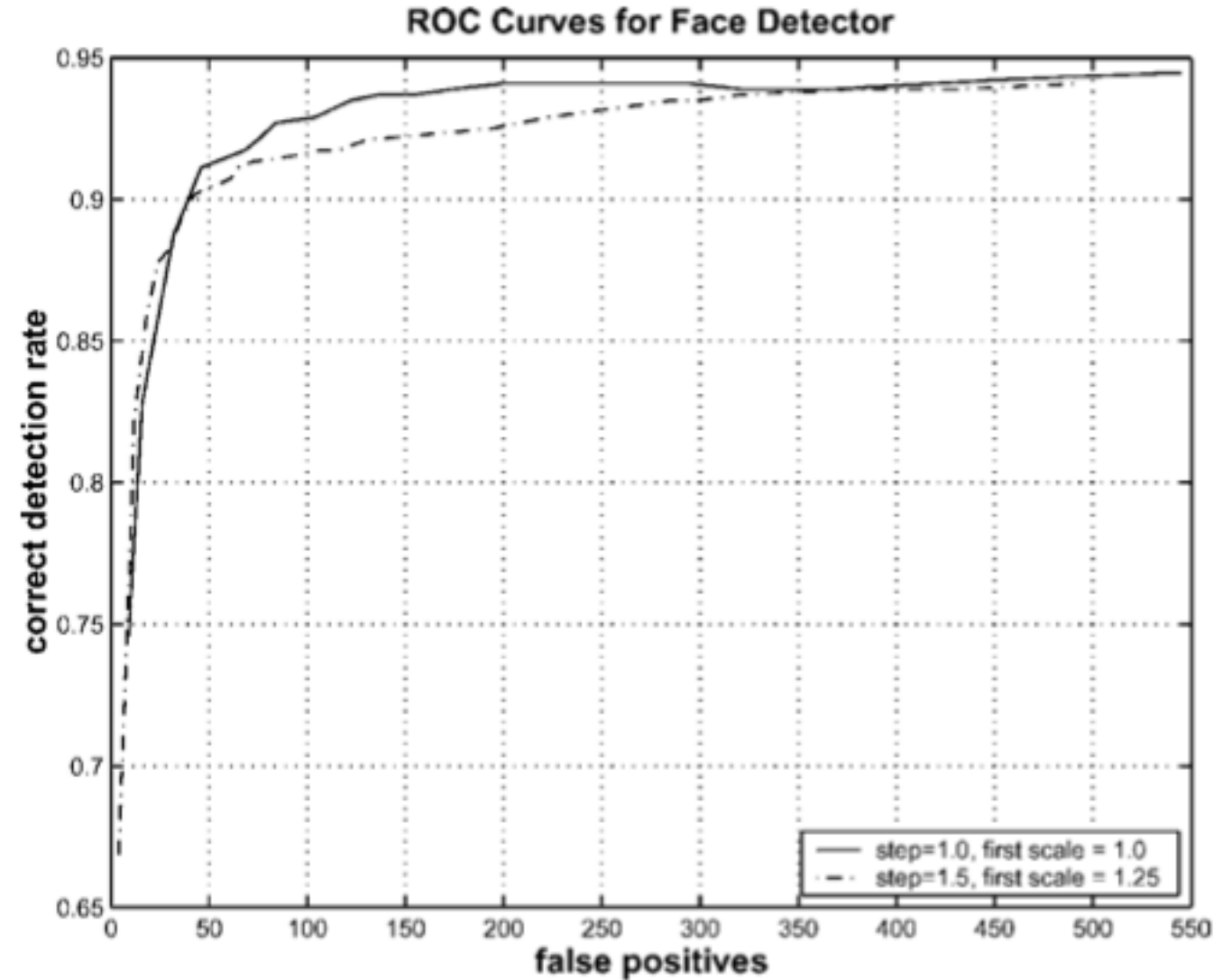
O **terceiro e o quarto classificador** utilizaram 25 *features* cada;

O **quinto, sexto e sétimo classificador** utilizaram 50 *features* cada;

O processo de escolha do numero de *features* por camada foi escolhido através de intervenção humana nas sete primeiras camadas, a fim de reduzir o tempo de processamento. Os demais foram selecionados através do algoritmo.

CASCADE

O detector de face pode processar uma imagem 384x288 pixel em 0.67 segundos utilizando um processador 700Mhz Pentium III;



CONSIDERAÇÕES

Não detecta faces inclinadas acima de 15 graus;

Não detecta faces quando o rosto está muito escuros e o fundo relativamente claro;

Não detecta faces se os olhos estiverem fechados;

A boca não é importante para a detecção de face.