

# Visão Computacional

## Aula 16

Reconhecimento de Padrões

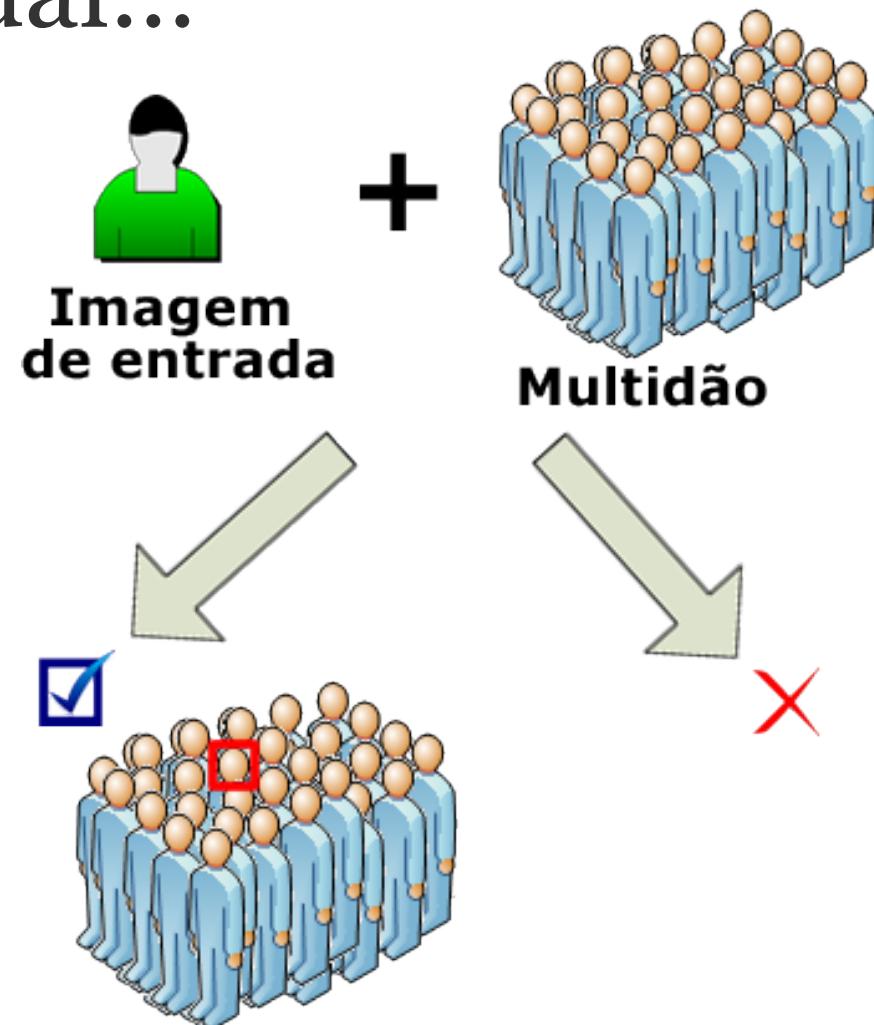
*Detecção e Reconhecimento Facial*

# Motivação

- Aplicações
- Reconhecimento de Faces a partir de Vídeo
  - Detecção e Rastreamento de Faces / Características Faciais
  - Segmentação e Normalização
  - Extração de Características e Reconhecimento
- Restrições e Definição do Problema

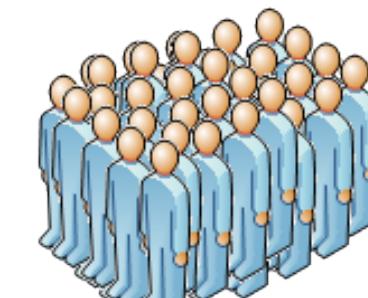
# O problema mais atual...

- Reconhecimento de faces em imagens que possuem diversas faces.
- Enunciado do problema:
  - **Opção 1:** Dada uma imagem de uma única face, encontre em uma multidão esta face.
  - **Opção 2:** Dada a imagem de uma multidão, reconheça as faces presentes no banco de dados da aplicação.

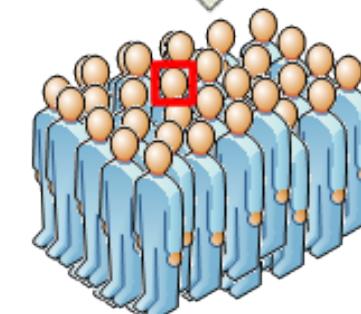
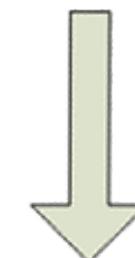


# E ainda...

- Identificação de criminosos, terroristas, mafiosos, baderneiros, etc em lugares públicos como Estações de Trem, Aeroportos, Feiras, Estádios.
  - Pode ajudar a prevenir crimes contra a vida em escala catastrófica como o "11/09".
- Motivação principal: **Segurança!**



Multidão



Criminoso

# É um problema simples?



- Há dificuldades com:
  - Pose
  - Qualidade da imagem
  - Oclusão
  - Disfarces
  - Desempenho

# Literatura

- PCA (Moghaddam e Pentland, 1997)
- Redes Neurais (Sung e Poggio, 1998; Rowley et al., 1998)
- Modelos de Cor (Yang e Waibel, 1996; Raja et al., 1998)
- Filtros de Gabor (Maurer e Malsburg, 1996)
- Modelos Paramétricos (Hager e Belhumeur, 1998)

# Literatura

- Rastreamento 3D da Face (Sclaroff et al., 2000)
- Detecção e Rastreamento de Características Faciais
  - (Feris et al., 2000; Silva et al., 1995; Morimoto et al., 1998)

# Como o problema é tratado

- Duas etapas:
  - Detecção da Face
    - Região da imagem digital que compõe/delimita a face;
  - Reconhecimento
    - Dado uma face de referência deve-se determinar de quem é a face.

# Detecção de Faces

- Principais trabalhos:
  - VIOLA, Paul; JONES, Michael (2003, 2004,...)
  - Robust Real-Time Face Detection (1<sup>a</sup> Versão)
  - Fast Multi-view Face Detection (2<sup>a</sup> Versão)
    - Considerado o estado-da-arte em detecção
    - Detecta faces com rotação dentro e fora do plano de imagem
    - Mantém o bom desempenho da primeira versão
    - 0,12 segundos para uma imagem 320x240 em um P4 2.8 GHz



# Reconhecimento de Faces

- Existem muitas técnicas...
  - Características faciais (olhos, nariz, boca)
    - Brunelli, R.; Poggio, T.; Trento, I Povo. *Face recognition through geometrical features.* 1992
  - Template Matching
  - Análise da Textura da pele
  - Principal Component Analysis - Eigenfaces
    - SIROVICH, L.; KIRBY, M.; *Low-dimensional Procedure for the characterization of human faces.* 1987
    - M. Turk, A. Pentland, *Eigenfaces for Recognition.* 1991 (+ citado)<sup>1</sup>
    - ATALAY, Ilker; *Face recognition using eigenfaces*

1) Face Recognition Homepage - Interesting Papers <http://www.face-rec.org/interesting-papers/>  
2) <http://electronics.howstuffworks.com/facialrecognition.htm>

# Reconhecimento de Faces

- Reconhecimento de Faces - Imagens Estáticas
  - Reconhecimento Tridimensional
    - Cartoux, J. Y., J. T. LaPrete, and M. Richetin: 1989, *Face authentication or recognition by profile extraction from range images*
    - Bronstein, A. M.; Bronstein, M. M., Kimmel, R. *Three-dimensional face recognition*. 2005
      - Identificou corretamente 2 gêmeos idênticos (autores)
    - Queirolo, C. C.; Silva, L.; Bellon, O. R.; Segundo, M. P. *3D Face Recognition using Simulated Annealing and the Surface Interpenetration Measure*. 2009
      - Precisão de identificação maior que 98%.

# Reconhecimento de Faces

- Reconhecimento de Faces em Multidões:
  - No entanto, alguns sistemas conhecidos:
    - FaceIt from Visionics
      - *Can find human faces anywhere in a field of view. It can track up to 10 faces simultaneously in a live video. It can follow that face as it moves through a crowd while also searching for matches against a database at a rate of 60 million/minute, per central processing unit (CPU).*
    - FaceVACS-Alert. <http://www.cognitec-systems.de/FaceVACS-Alert.20.0.html>
      - *Real time face tracking on multiple video streams*
      - *Real time probe face comparison against "watch list"*
      - *Real time alarm notification and recordind*
    - Biometric Systems, Inc. <http://www.biometrica.com/>
      - *providing advanced identity solutions to the Casino Industry*

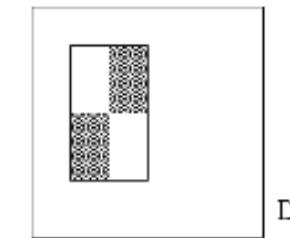
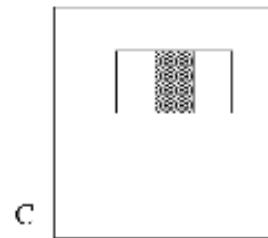
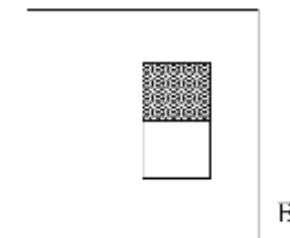
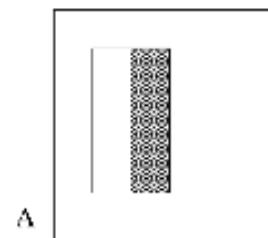
# Detectando Faces

- Método de Viola-Jones (2003)
  - Utilizado para detecção em tempo real;
  - Treinamento lento, mas a detecção é muito rápida!
  - Idéias Básicas:
    - Utiliza imagens integrais para extração de característica rápida da face;
    - “Boosting” para seleção da característica da face;
    - Técnica de Cascata (cascade)
      - Várias escalas diferentes;

# Detectando Faces

- Método de Viola-Jones (2003)
  - Característica da Imagem
    - Filtros Retangulares

“Rectangle filters”

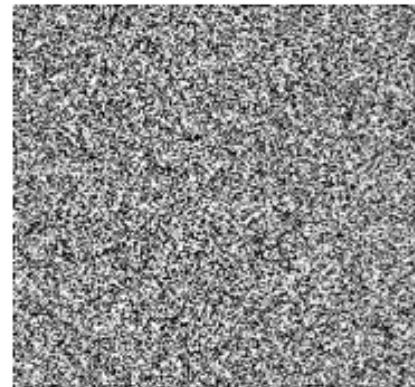


*Value =*

$$\sum (\text{pixels in white area}) - \sum (\text{pixels in black area})$$

# Detectando Faces

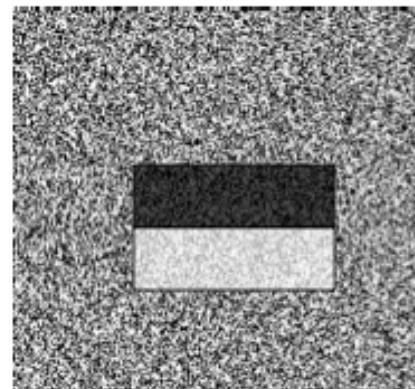
- Método de Viola-Jones (2003)
  - Característica da Imagem
  - Exemplo



Source

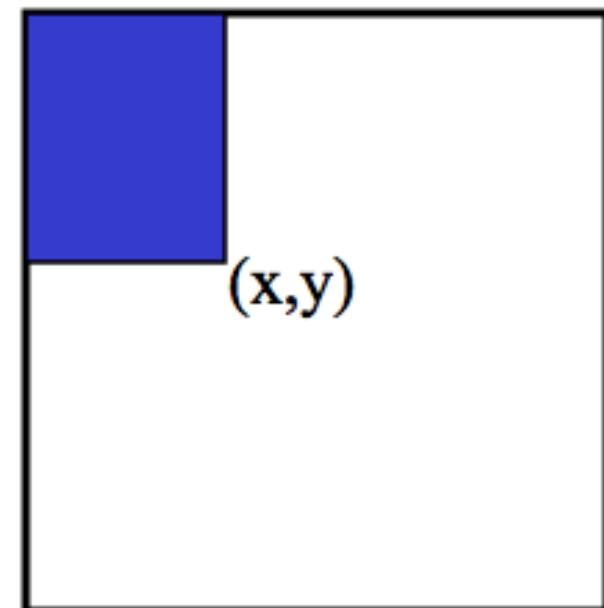


Result



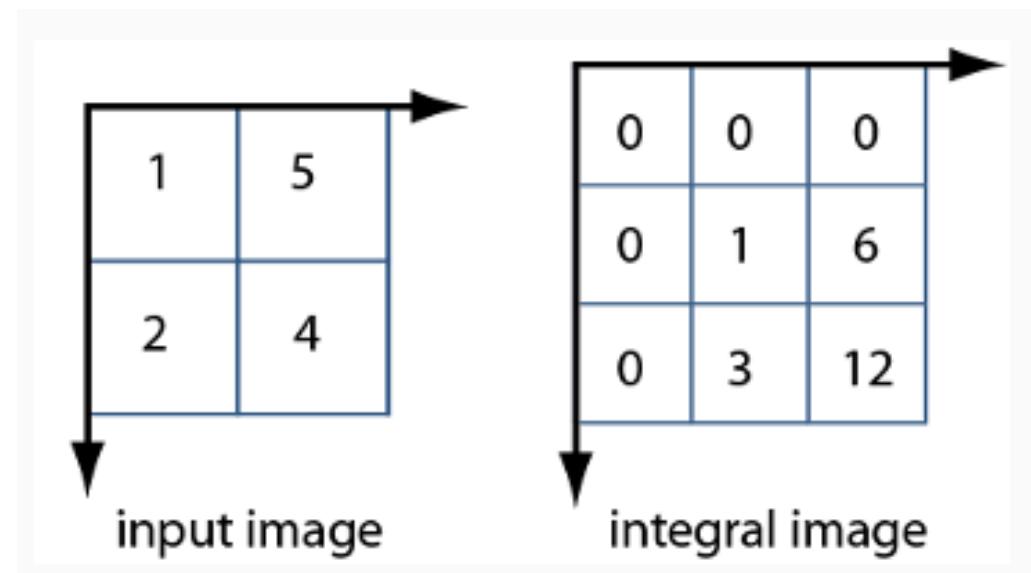
# Detectando Faces

- Método de Viola-Jones (2003)
  - Imagem Integral
    - Método rápido e utilizando uma única passagem pela imagem de entrada;
    - É obtida para cada pixel  $(x,y)$  que é a soma dos valores dos pixels acima e à esquerda de  $(x,y)$ , incluindo o ponto  $(x,y)$ .



# Detectando Faces

- Método de Viola-Jones (2003)
  - Imagem Integral



# Detectando Faces

- Método de Viola-Jones (2003)
  - Imagem Integral

1	2	2	4	1
3	4	1	5	2
2	3	3	2	4
4	1	5	4	6
6	3	2	1	3

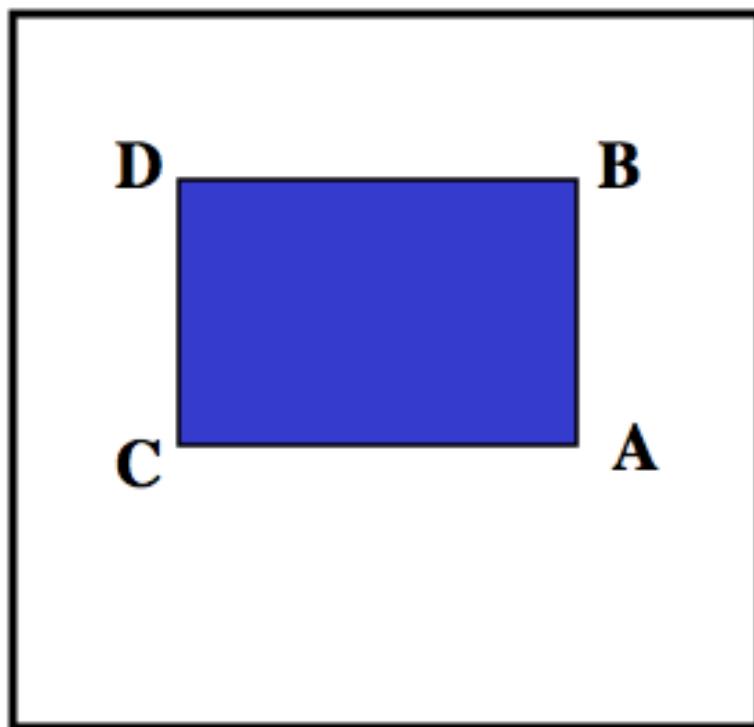
input image

0	0	0	0	0	0
0	1	3	5	9	10
0	4	10	13	22	25
0	6	15	21	32	39
0	10	20	31	46	59
0	16	29	42	58	74

integral image

# Detectando Faces

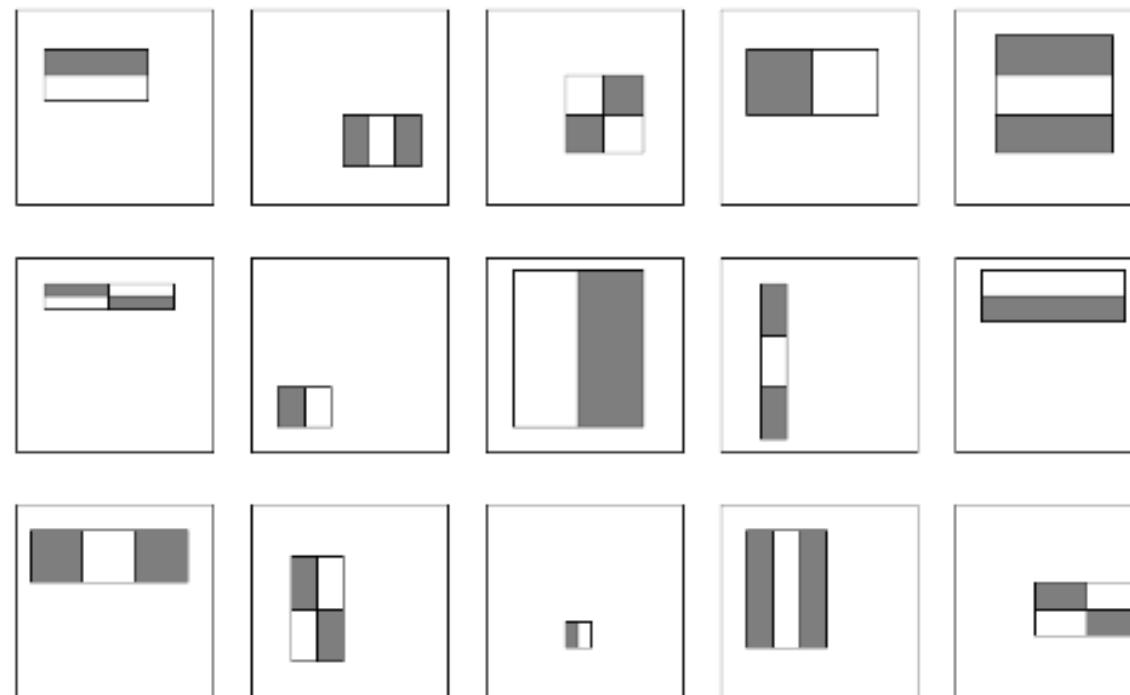
- Método de Viola-Jones (2003)
  - Imagem Integral
    - A partir da imagem integral, pode se calcular a soma da área da região de interesse da imagem facilmente, por exemplo:



$$\text{sum} = A - B - C + D$$

# Detectando Faces

- Método de Viola-Jones (2003)
  - Extração da característica da Face
    - Para uma região de detecção de 24 x 24 pixels o número de retângulos pode chegar a 180.000 possíveis combinações!



# Detectando Faces

- Método de Viola-Jones (2003)
  - Extração da característica da Face
    - Para solucionar o problema, utilizou-se uma técnica de “Boosting” – AdaBoosting – para “melhorar classificadores fracos”.

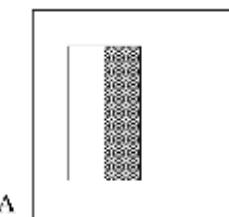
“Rectangle filters”



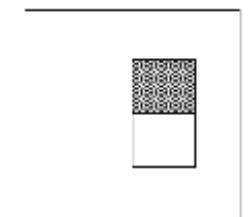
Similar to Haar wavelets

Papageorgiou, et al.

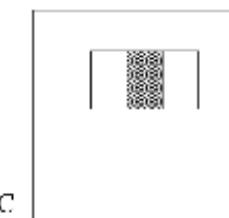
$$h_t(x_i) = \begin{cases} \alpha_t & \text{if } f_t(x_i) > \theta_t \\ \beta_t & \text{otherwise} \end{cases}$$



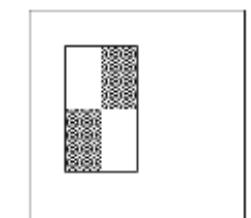
A



B



C



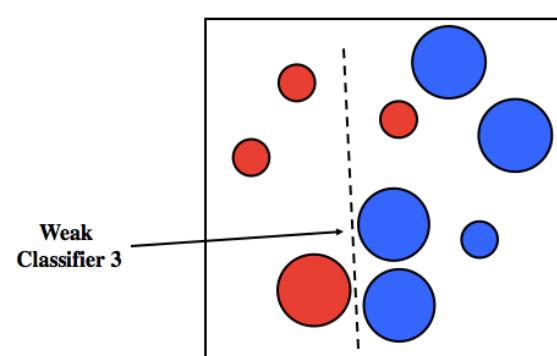
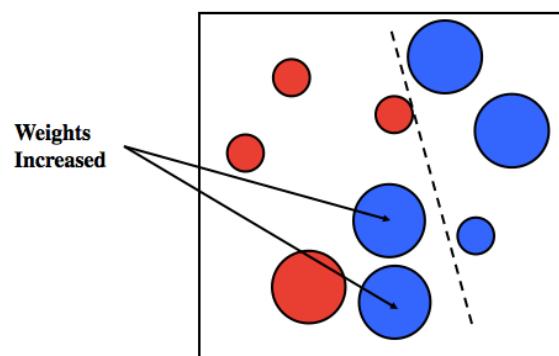
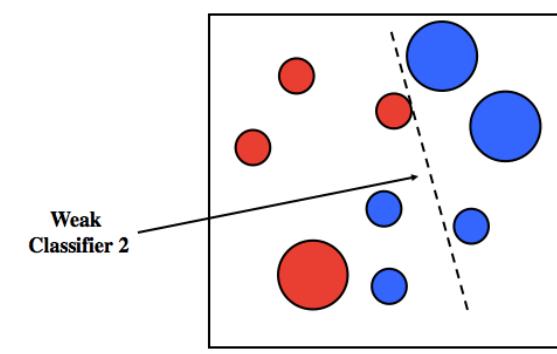
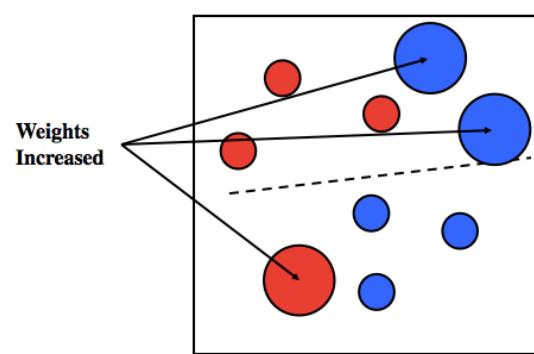
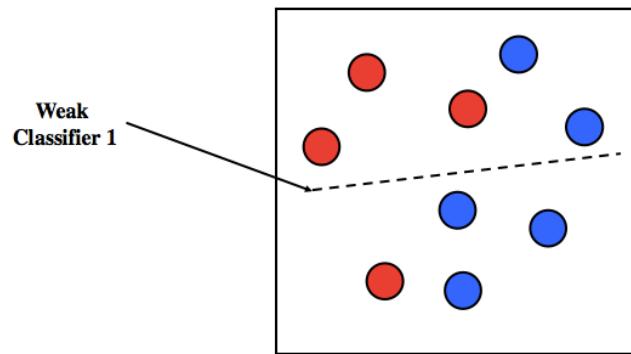
D

$$C(x) = \theta \left( \sum_t h_t(x) + b \right)$$

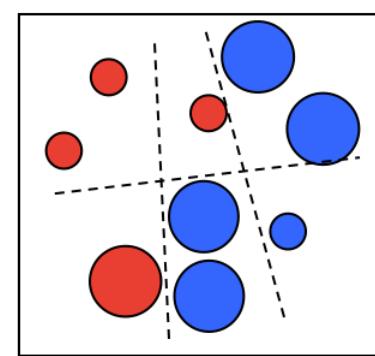
60,000 features to choose from

# Detectando Faces

- Método de Viola-Jones (2003)
  - Extração da característica da Face
    - Classificadores Fracos x Boosting

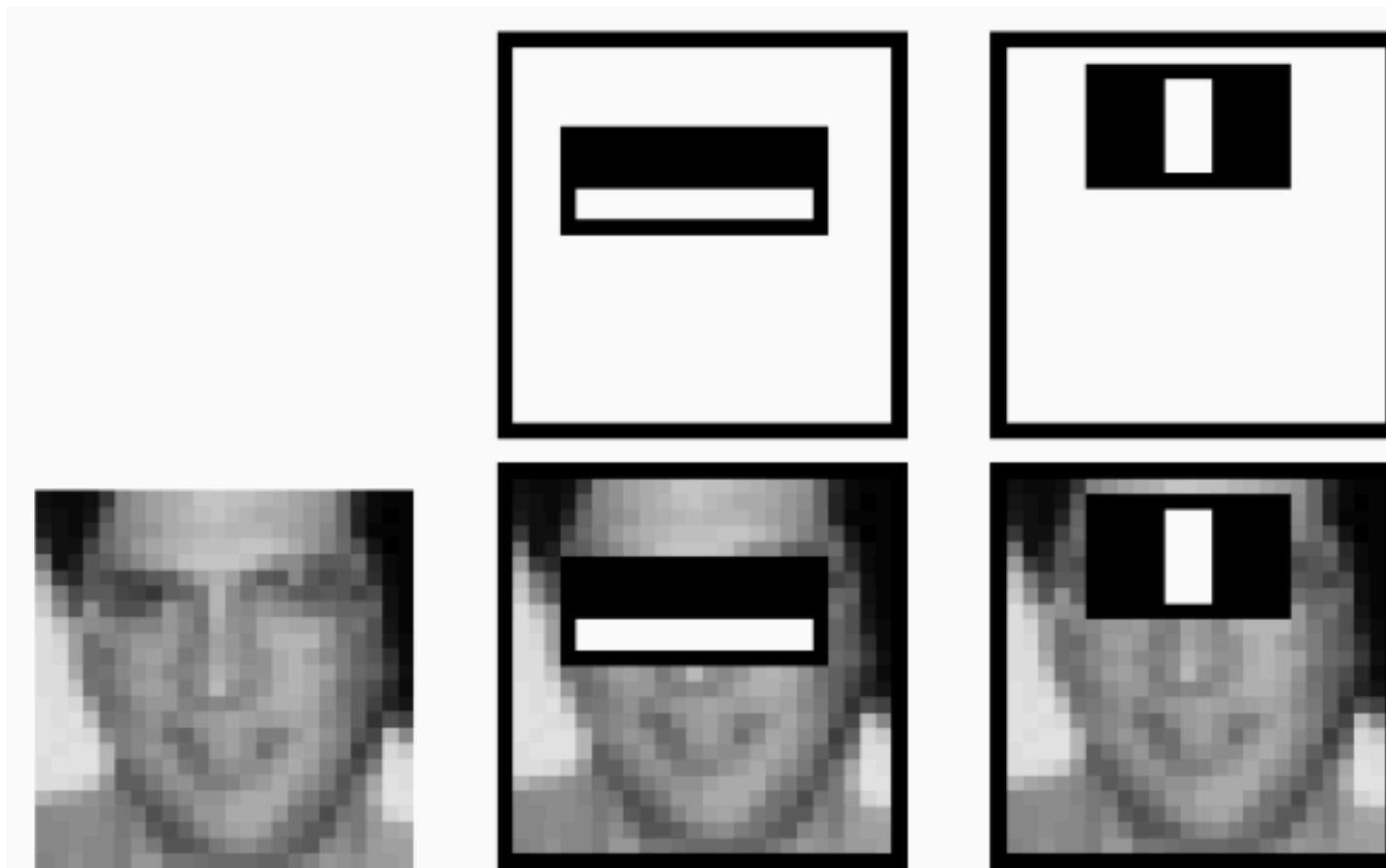


Final classifier is  
linear combination of  
weak classifiers



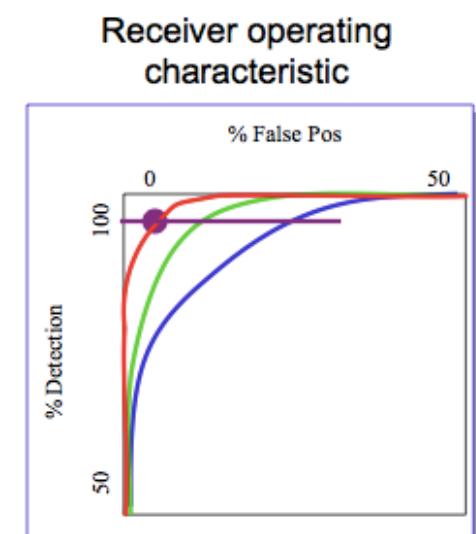
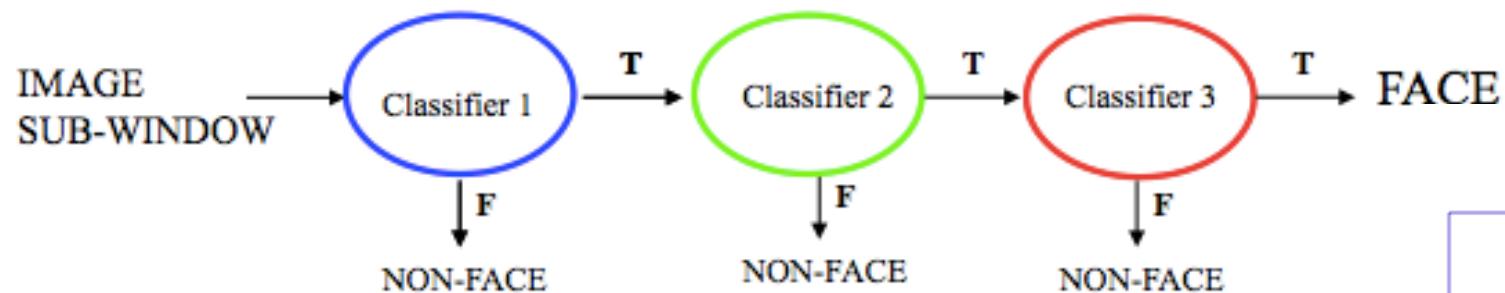
# Detectando Faces

- Método de Viola-Jones (2003)
  - Extração da característica da Face
    - Usando duas features c/ boosting



# Detectando Faces

- Método de Viola-Jones (2003)
  - Classificadores em Cascata
    - Inicialmente emprega-se classificadores simples para avaliar a sub-imagem (ROI);
    - Resultados positivos são propagados ao próximo classificador e negativos imediatamente rejeitados.



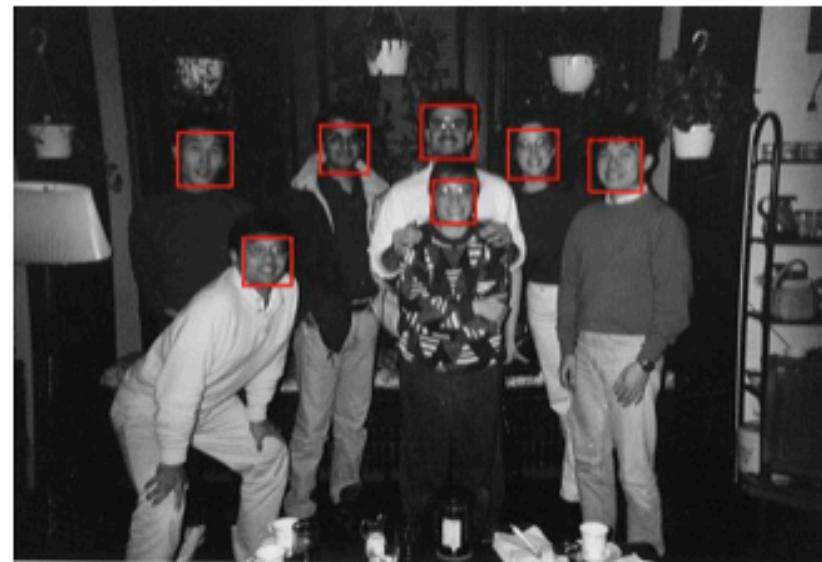
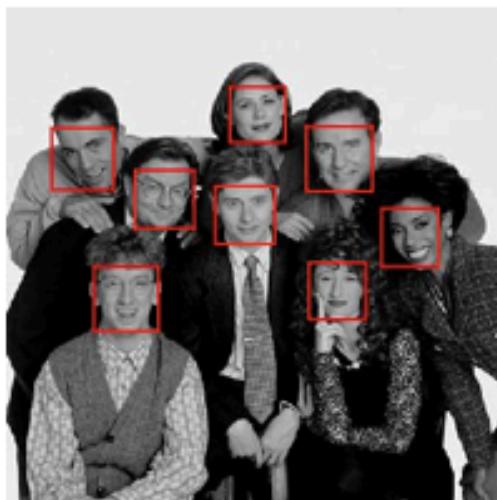
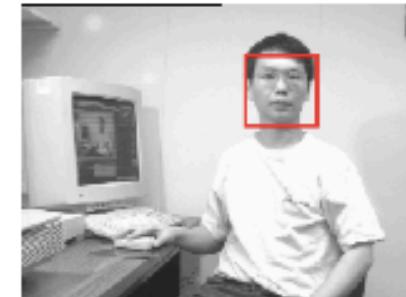
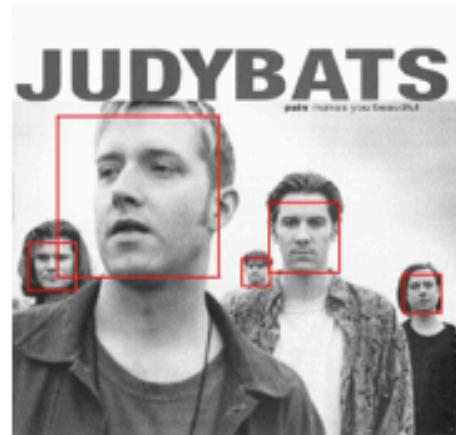
# Detectando Faces

- Método de Viola-Jones (2003)
  - Sistema Desenvolvido por Viola-Jones
- **Training Data**
  - 5000 faces
    - All frontal, rescaled to 24x24 pixels
  - 300 million non-faces
    - 9500 non-face images
  - Faces are normalized
    - Scale, translation
- **Many variations**
  - Across individuals
  - Illumination
  - Pose



# Detectando Faces

- Método de Viola-Jones (2003)
  - Resultados obtidos:





# Após a detecção...

- Realizar o Reconhecimento Facial!

## Reconhecimento Facial na Psicologia

- Níveis do reconhecimento da face:
  - Reconhecimento em nível de entrada; e
  - Reconhecimento do em nível subordinado
- O cérebro tem regiões específicas para o Reconhecimento Facial

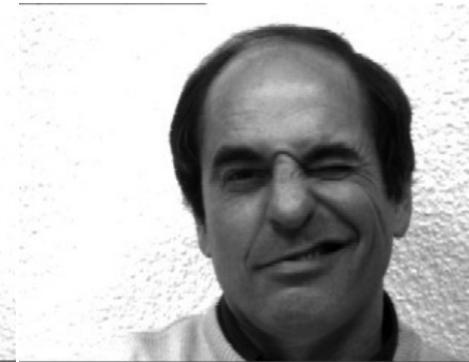
## Etapas do Reconhecimento Facial

- Representação Facial
  - *Template-based*
  - *Feature-based*
  - *Appearance-based*

# Fatores que interferem no desempenho do RF

- Expressões Faciais
- Iluminação inadequada
- Disfarces
- Escala
- Posição da Face

# Expressões Faciais



## Iluminação inadequada



# Disfarces



# Escala



# Posição da Face



# Principais Técnicas

- Análise de Componentes Principais
- Classificadores Intra Bayesianos - BIC
- Técnicas Alternativas:
  - Redes Neurais Artificiais
  - Hidden Markov Models
  - *Eigenfaces*

# Principais Técnicas

- **Principal Component Analysis (PCA)<sup>1</sup>**
  - Forma vetores de características (feature vectors) concatenando valores de pixels da imagem
  - Inicialmente são gerados muitos valores altamente correlacionados
  - PCA chega a um subespaço pequeno sem covariância entre os *feature vectors* transformados (os componentes principais).
  - A matriz de covariância é sempre diagonal
    - Isso é explorado para alguns classificadores, como L1, *MahCosine* e Bayesianos, por exemplo.

<sup>1</sup> - M. A. Turk and A. P. Pentland. Face Recognition Using Eigenfaces. In Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pages 586 – 591, June 1991.

# Principais Técnicas

- **PCA - Métricas de Classificação**

- **Euclideana:**  $D_{Euclidean}(u, v) = \sqrt{\sum_i (u_i - v_i)^2}$

- **Mahalanobis Distance**

- Imagem é convertida para o espaço de Mahalanobis
- A variância da *amostra* ao longo de cada dimensão é unitária, contra o dos feature vectors
- Transformação é feita pela divisão dos componentes do vetor pelos seus respectivos desvios-padrão.

Transformação:      MahCosine:       $D_{MahCosine}(u, v) = -S_{MahCosine}(u, v)$

$$m_i = \frac{u_i}{\sigma_i}$$

$$S_{MahCosine}(u, v) = \cos(\theta_{mn}) = \frac{|m| |n| \cos(\theta_{mn})}{|m| |n|} = \frac{m \cdot n}{|m| |n|}$$

(covariância entre as imagens no espaço de Mahalanobis)

# Principais Técnicas

- **Bayesian Intra/Extra personal Classifier (BIC)<sup>1</sup>**
  - Examina o resultado da diferença entre imagens
  - As imagens-diferença são originadas de distribuições Gaussianas dentro do espaço de todas as imagens-diferença
  - O treinamento é feito através de PCA, determinando as propriedades estatísticas de 2 subespaços
    - Imagens-diferença da classe Intrapessoal
    - Imagens-diferença da classe Extrapessoal
  - Na fase de teste, o classificador recebe uma imagem de classe desconhecida e usa as estimativas das distribuições de probabilidade para identificação.
    - Na prática: os *feature vectors* são projetados nos dois conjuntos, e a probabilidade dele pertencer a cada subespaço é calculada.

<sup>1</sup> - B. Moghaddam, C. Nastar, and A. Pentland. A bayesian similarity measure for direct image matching. ICPR, B:350–358, 1996.

# Principais Técnicas

- **BIC - Métricas de Classificação**

*Os feature vectors são projetados nos dois conjuntos, e a probabilidade dele pertencer a cada subespaço é calculada.*

- **MAP - Maximum a posteriori**
  - Métrica gerada em relação ao conjunto de imagens intra e extrapessoais.
- **ML - Maximum likelihood**
  - Gerada em relação apenas ao conjunto de imagens intrapessoais.

# Banco de Dados de Faces utilizado para avaliação

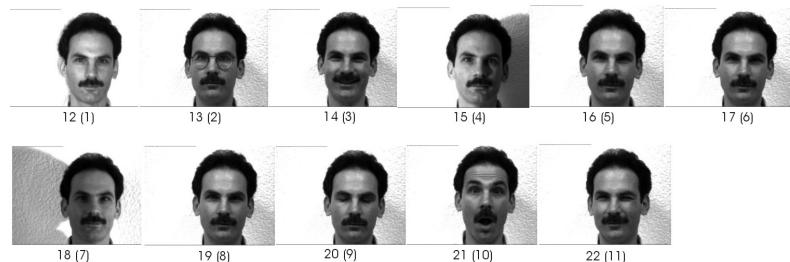
- The Yale Face DataBase
- 15 classes
- 11 imagens p/classe: **centerlight**, withglasses, happy, **leftlight**, noglasses, normal, **rightlight**, sad, sleepy, surprised e wink/blink

# Banco de Dados de Faces

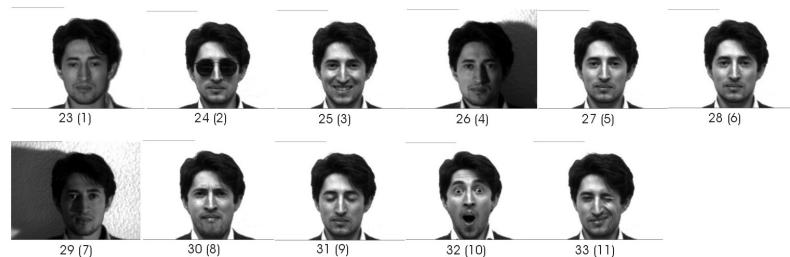
CLASSE 01



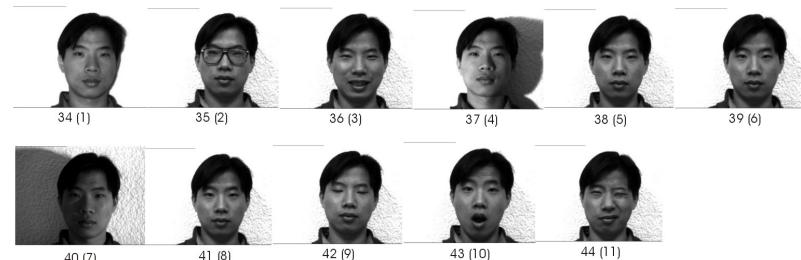
CLASSE 02



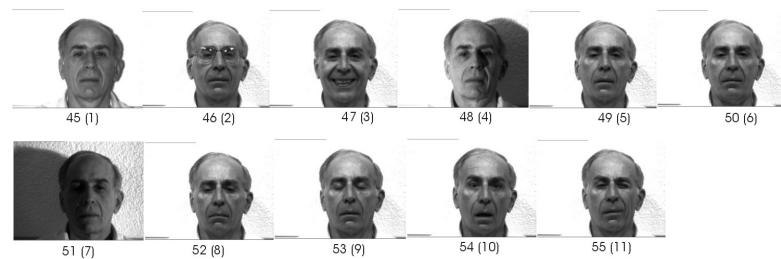
CLASSE 03



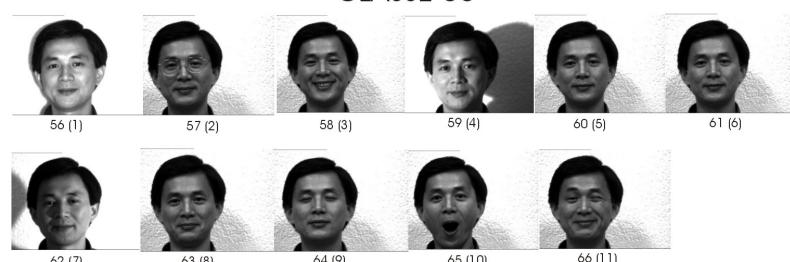
CLASSE 04



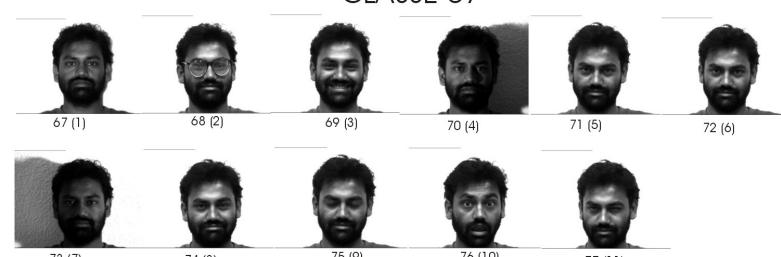
CLASSE 05



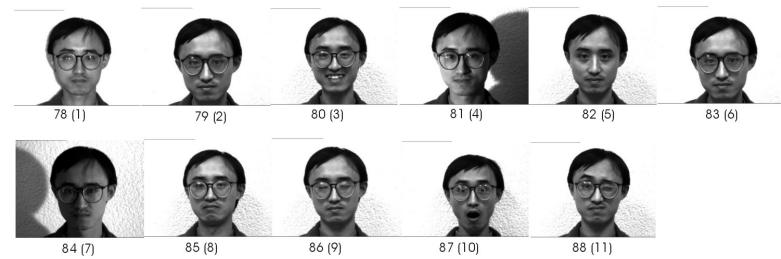
CLASSE 06



CLASSE 07



CLASSE 08



# Próxima aula...

- Movimento em Imagens