

# Processamento digital de imagens

Agostinho Brito

Departamento de Engenharia da Computação e Automação  
Universidade Federal do Rio Grande do Norte

1 de dezembro de 2016

# Algoritmo de Viola-Jones

- Proposto em 2001 por Paul Viola e Michael Jones
- Amplamente usado para detecção em tempo real de objetos
- Treinamento lento, mas de rápida detecção.
- Operando em um **Pentium III de 700MHz**, o algoritmo foi capaz de processar até 15 fps em imagens de 384 por 288 pixels.

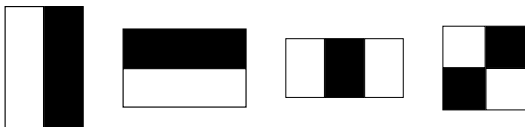
# Como achar a face?



# Como achar a face?

- Em cada sub-janela o algoritmo precisa decidir se há uma face lá dentro.
- Redes neurais podem ajudar.... :-/
- Problemas:
  - A solução precisa operar em tempo real.
  - Iluminação.
  - Indivíduos diferentes
  - Posicionamento da face
  - Faces são raras por imagens
- Proposta: usar combinações de classificadores fracos (e simples) para realizar a detecção da face.

- Tipos de características



- Medida da característica: soma dos pixels nos quadrados pretos menos a soma dos pixels nos quadrados brancos.

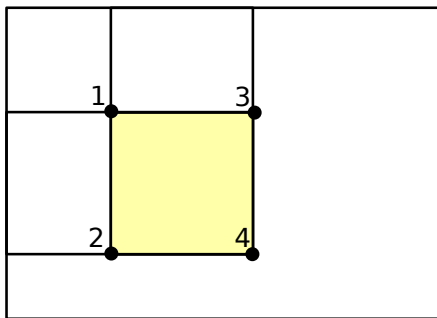
$$caracteristica = \sum \blacksquare - \sum \square$$

- Problema: muitas combinações para um detector com resolução de  $24 \times 24$  pixels.
- Solução: encontrar as melhores características. Realizar cálculos rápidos. Uso de Adaboost e imagem integral.

# Imagem integral

- A imagem integral de uma imagem é construída com soma de todos os pixels acima e à esquerda de um dado ponto.

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



- Dada por

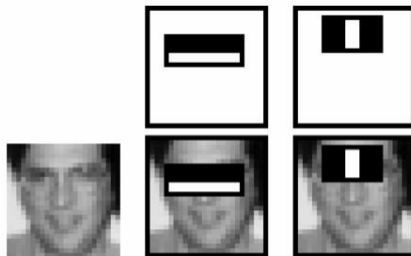
$$II = \sum (ii_4 + ii_1 - ii_2 - ii_3)$$

- Um classificador fraco é dado por

$$h_j(x) = \begin{cases} 1 & \text{se } p_j f_j(x) < p_j \theta_j \\ 0 & \text{caso contrário} \end{cases}$$

- Na equação,  $x$  é uma janela de  $24 \times 24$  pixels.  $\theta$  é um threshold.  $p_j$  é uma polarização (+1 ou -1)
- Na escolha dos classificadores, a característica mais apropriada é seleccionada

- Teste todos os filtros retângulos possíveis em múltiplas rodadas.
- Para cada rodada
  - Avalie todos os filtros
  - Selecione o melhor threshold para cada filtro
  - Procure a melhor combinação filtro/threshold.
  - Aumente os pesos dos exemplos com classificação errada.
  - São  $T$  rodadas,  $N$  exemplos e  $K$  características. Complexidade:  $O(T \times N \times K)$ .
- Exemplo de características selecionadas







## Aspectos de implementação

- Todas as imagens são normalizadas para evitar problemas de iluminação.
- O detector desenvolvido por Viola e Jones tem 38 níveis de cascata e 6000 características.
- Em um conjunto de dados de 507 faces, e 75 milhões de subjanelas, uma média de 10 avaliações de características é feita para cada subjanela.
- A redução da escala do detector é feita redimensionando o detector e não a própria imagem. Assim, as características podem ser calculadas em qualquer escala com o mesmo custo computacional.
- As cinco primeiras camadas possuíam 1, 10, 25, 25 e 50 características, respectivamente.