Haar Cascades

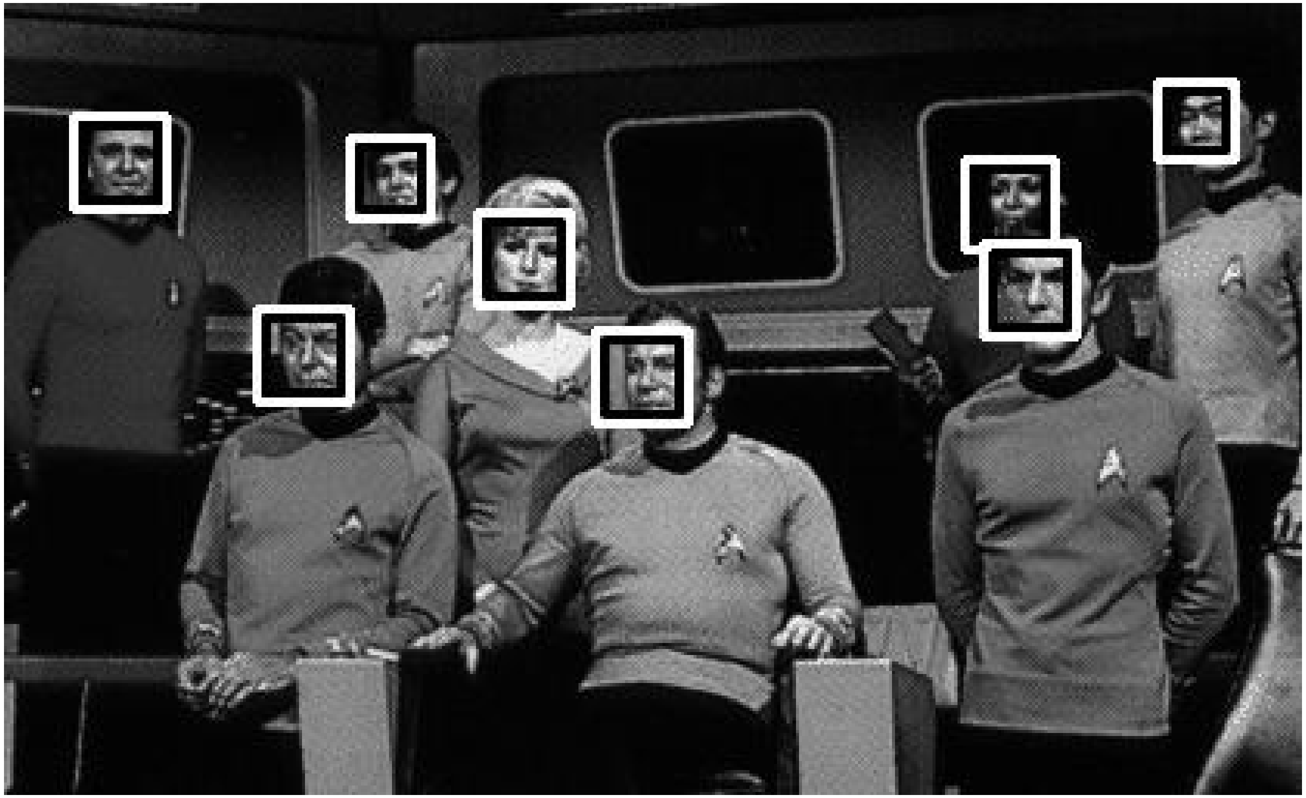
BM40A0800
MACHINE VISION AND DIGITAL IMAGE ANALYSIS

Authors:
Tikhon Belousko
Iakov Lushin

Plan

- Introduction
- Theoretical background
- Application demonstration
- Comparison with other methods
- Conclusion

Introduction



Face detection is a general application of the Haar Cascades.

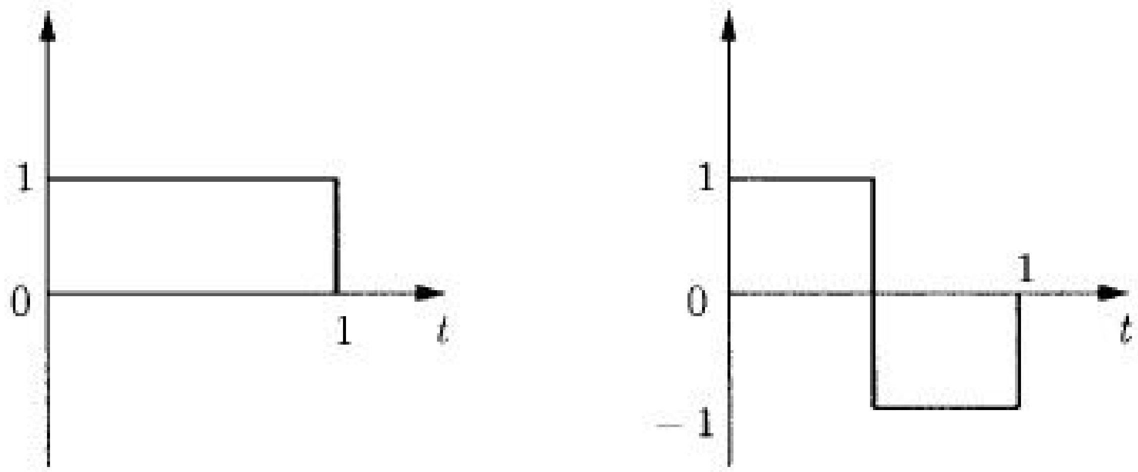
Theoretical Background

Object detection process is usually performed by scanning the original image with a detector.

Theoretical Background: Main Concepts

- Haar-like features
- Integral images
- AdaBoost
- Attentional cascades

Theoretical Background: Haar-like Features

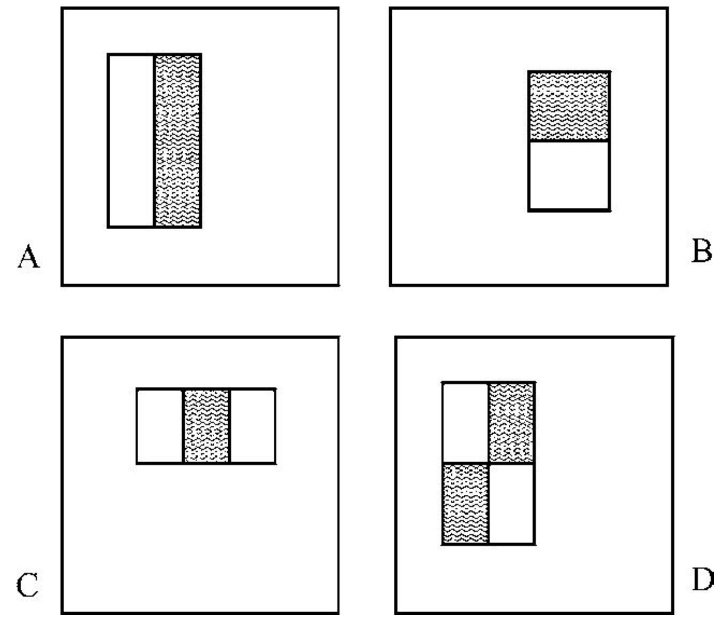


(a) (b)

The Haar Basis: a) Scaling function; b) Wavelet [1].

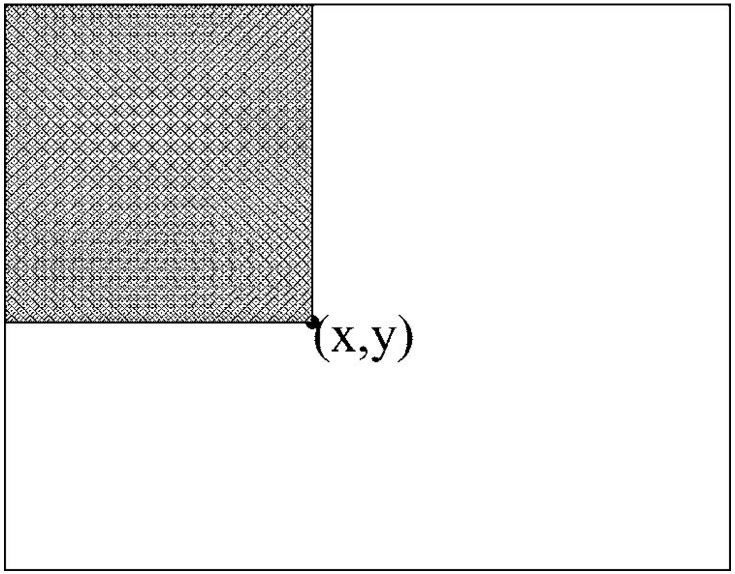
Theoretical Background: Haar-like Features

- Figures A and B:
  two-rectangle feature.
- Figure C:
  three-rectangle feature.
- Figure D:
  four-rectangle feature.



Example rectangle features shown relative to the enclosing detection window [2].

Theoretical Background: Integral Images



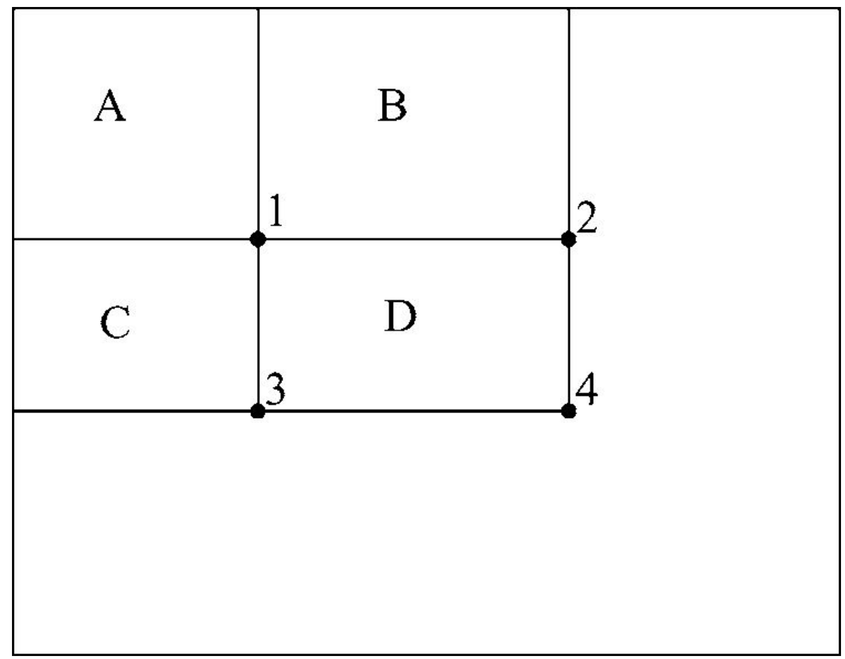
The value of the integral image at point (x, y) is the sum of all the pixels above and to the left [2].

Theoretical Background: Integral Images

- X1 = A
- X2 = A + B
- X3 = A + C
- X4 = A + B + C + D

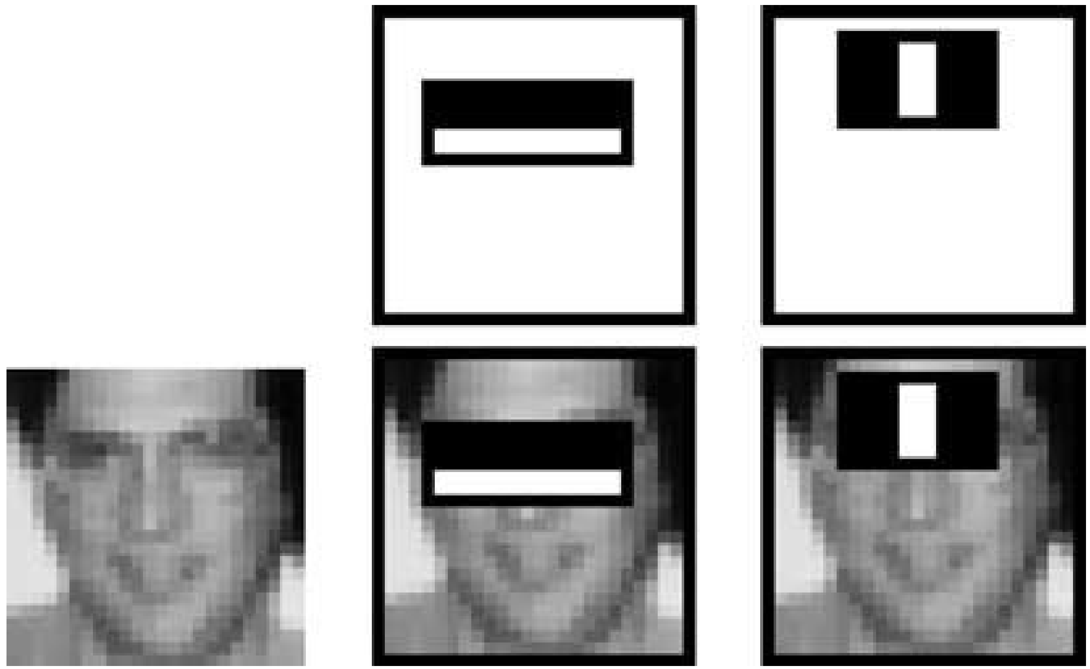
Then

- D = (X4 + X1) - (X3 + X2)



Computation of the sum of the pixels within rectangle D [2].

Theoretical Background: AdaBoost Classifiers



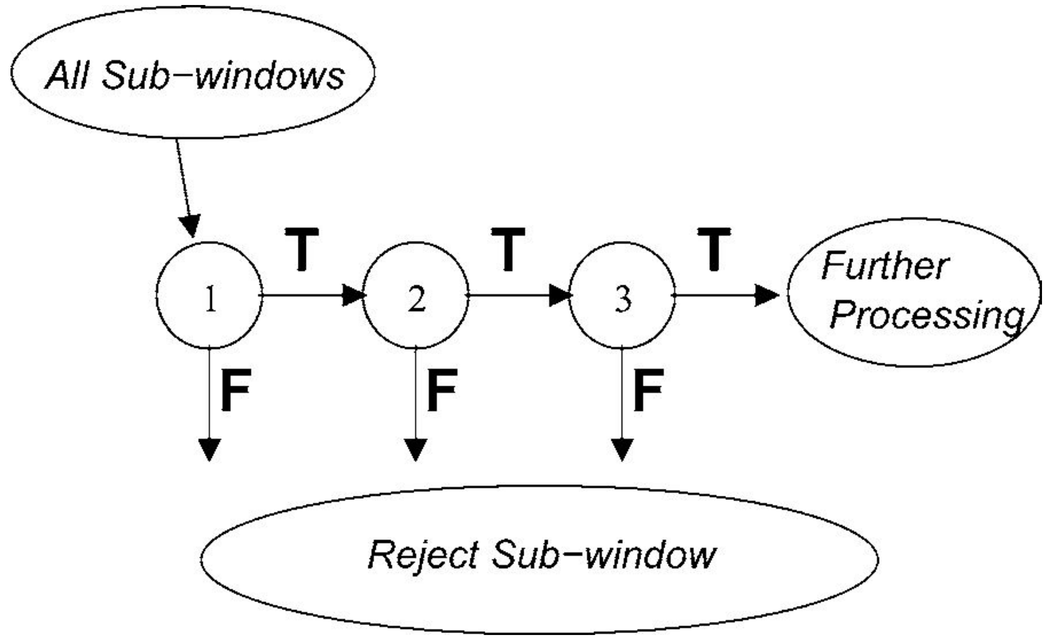
- Strong classifier constructed from the set of weak classifiers.
- Weighted sum of binary classifiers (one Haar-like feature for each classifier).
- Each feature represents changes of intensities.

Theoretical Background: Attentional Cascades

Problem: Evaluation of strong classifiers with many features is slow.

Solution: Use cascade of classifiers starting from simple one which will reject most of the false positives and finishing with strong classifier which will need to analyze smaller amount of subwindows after cascading.

Theoretical Background: Attentional Cascades



Deeper layer:

- Slower
- More precise
- Analyses less subwindows

Theoretical Background: Training process

- Each layer aims to discard as many subwindows as possible, while keeping all or most of the subwindows, that really contain the objects.
- Get the number of false negatives close to zero in addition to just lowering the error rate.

Theoretical Background: Training process

- **Layer 1:**
  - Positive samples
  - Negative samples
- **Layer 2:**
  - Positive samples
  - False Positives from Layer 1

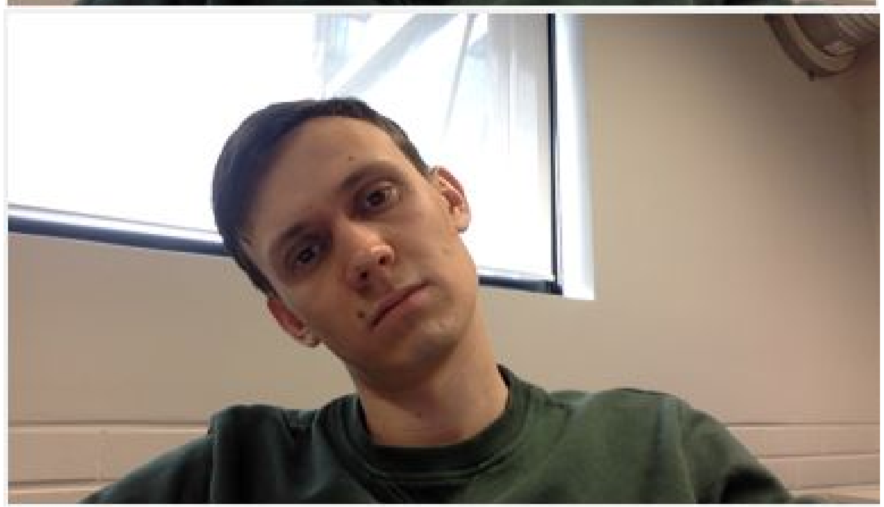
...

- **Layer N:**
  - Positive samples
  - False Positives from Layers 1 to N-1

Theoretical Background: Training process

- The number of features on each layer depend on system requirements
- The features are added until certain rate of false positives is reached
- The layers are added until desired overall effectiveness is reached

Demonstration: Real-Time Face Detection



Demonstration: Road Sign Detection

Comparison to Other Methods

Extensions of Viola-Jones Algorithm:

- Rotate rectangles by 45º [3]
- Center-surround features [3]
- More rectangle rotations [4]

Other Face Detection Methods:

- Soft cascades [5]
- Usage of SURF-based features [6]
- AUC learning criterion [6]
- Local Gradient Patterns [7]
- Evidence Accumulation Method [7]

Conclusion

- The face detection algorithm by Viola and Jones is described
- The main steps of the algorithm are explained
- The alternatives to the Haar Cascades are introduced

References

[1] P. Kechichian, “On the Partial Haar Dual Adaptive Filter for Sparse Echo Cancellation,” M.E. thesis, McGill University, Canada, 2006.

[2] P. Viola and M. J. Jones, “Robust Real-Time Face Detection,” International Journal of Computer Vision, vol. 57, no. 2, pp. 137–154, May 2004.

[3] R. Lienhart and J. Maydt, “An extended set of Haar-like features for rapid object detection,” in 2002 International Conference on Image Processing. 2002. Proceedings, 2002, vol. 1, pp. I–900–I–903 vol.1.

[4] C. Messom and A. Barczak, “Fast and efficient rotated haar-like features using rotated integral images,” presented at the Proceedings of the 2006 Australasian Conference on Robotics and Automation, ACRA 2006, 2006.

References

[5] L. Bourdev and J. Brandt, “Robust object detection via soft cascade,” in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005, 2005, vol. 2, pp. 236–243 vol. 2.

[6] J. Li, T. Wang, and Y. Zhang, “Face detection using SURF cascade,” in 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), 2011, pp. 2183–2190.

[7] B. Jun and D. Kim, “Robust face detection using local gradient patterns and evidence accumulation,” Pattern Recognition, vol. 45, no. 9, pp. 3304–3316, Sep. 2012.