Report: Best Face Pose Detection

Roberto Capobianco and Jacopo Serafin

July 14, 2014

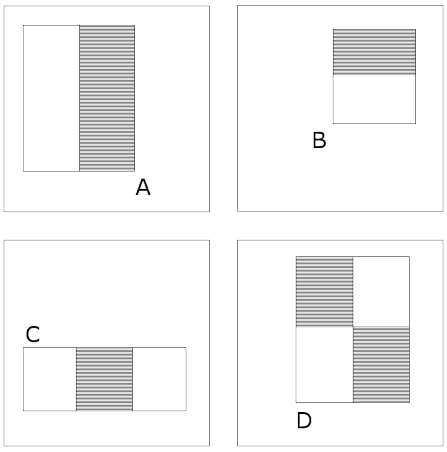

Figure 1: Haar-like features used in the Viola-Jones algorithm

Introduction and Problem Statement

Face detection is a particular subset of object-class detection. Differently from object-class detection, where the goal is to find the locations and sizes of all objects in an image that belong to a given class, face-detection algorithms focus on the detection of human faces.

Face detection is a fundamental building block of many applications, like:

Facial Recognition A computer application for automatically identifying a person from a digital image or a video frame. One possible way to do this is by comparing facial features between the image and a facial database.

Photography In recent digital cameras a basic features like the auto focus uses face detection to find regions in the image to focus.

Social Media Suggest to tag a friend in a photo is a basic peculiarity of almost all social network.

The most famous and commonly used approach to face detection is the Viola-Jones algorithm [1]. It is based on a set of features which involve the sums of image pixels within rectangular areas (integral images). Specifically, they are a more complex version of Haar basis functions, which have been used previously in image-based object detection [2], since they rely on more than one rectangular area. Figure 1 shows the types of features used in the framework presented by Viola and Jones. By using integral images, rectangular features can be evaluated in constant time, obtaining in this way a valuable speed-up over their more complex relatives, like steerable filters [3]. In many applications, however, being able to recognize frontal poses (or "best poses") in a video sequence or in a photo is a fundamental part since it improves drastically the results. For example, if we want to recognize a user, by comparing his face with a database of previously acquired faces, we need to have them in similar poses. Our aim is to construct a system which is able to better detect frontal poses of a single person in a video source. This is based on geometrical constraints on facial features.

Methodology

Our method is based on two fundamental steps: the first one is the face detection, along with the eyes and the mouth, in a video sequence; the second step is the worst pose rejection based on a series of geometric constraints related to the previously extracted facial features.

Feature Detection

As already stated, the first phase of the algorithm is to find, in real time, the face, eyes and mouth inside each frame of a video source stream. This is done exploiting the Viola-Jones method implementation of the OpenCV library[1]. More in detail we apply the following process steps:

1. Open a video stream and analyze each single frame
2. Face detection
3. Eyes detection
4. Mouth detection

Best Pose Selection

Results

Conclusions

[1] http://opencv.org/

Bibliography

[1] Viola, Paul, and Michael J. Jones. *"Robust real-time face detection."* International journal of computer vision 57.2 (2004): 137-154.

[2] Papageorgiou, Constantine P., Michael Oren, and Tomaso Poggio. *"A general framework for object detection."* Computer vision, 1998. sixth international conference on. IEEE, 1998.

[3] Freeman, William T., and Edward H. Adelson. *"The design and use of steerable filters."* IEEE Transactions on Pattern analysis and machine intelligence 13.9 (1991): 891-906.