LZM in Action:
Realtime Face Recognition System

Evangelos Sarıyanidi[1], Birkan Tunç[2], and Muhittin Gökmen[3]

[1] Control Engineering Department,
[2] Informatics Institute,
[3] Computer Engineering Department,
Istanbul Technical University

Abstract. In this technical demonstration, we introduce a real time face detection and recognition prototype. The proposed system can work with different image sources such as still images, videos from web cameras , and videos from ip cameras. The captured images are firstly processed by a cascaded classifier of Modified Census Transform (MCT) features to detect the faces. Then, facial features are detected inside the face region. These features are used to align and crop the face patches. Detection phase can be considerably improved by incorporating a tracking scheme to increase the hit rate while decreasing the false alarm rate. The registered faces are recognized using a novel method called Local Zernike Moments (LZM). A probabilistic decision step is employed in the final inference phase to provide a confidence margin. Introducing new identities via system's user interface is considerably simple since the system does not require retraining after each new identity.

1 Introduction

This demonstration introduces a real time face detection and recognition system which is developed to provide a solution to the face recognition problem against changing poses, illumination and facial expressions. The system consists of four interconnected blocks to perform (1) face detection, (2) face tracking, (3) face alignment and cropping, and (4) face recognition. The output of the final block is processed in a final probabilistic decision step to provide a recognition result along with a confidence margin.

The first block finds the faces present in a given image. The classifier used in this step is based on Modified Census Transform (MCT) [1]. The face scanning procedure is introduced in [2]. The detected face rectangles are used as the inputs of the second block to initiate a tracking process. Tracking provides a smooth transition between successive frames while decreasing the false alarm rate considerably by eliminating most of false alarms detected which are detected on single frames. The face alignment and cropping is performed in the third block by detecting facial features inside each face rectangle. MCT-based cascades are employed to detect facial features like eyes, mouth, and nose as they are used to detect faces. Each detected face is aligned according to a base patch and then cropped to provide a region of interest.

In the final block face recognition is performed using Local Zernike Moment (LZM) histograms [3]. Several moment components are calculated. These components are partitioned into subregions to get orientation-phase histograms inside each subregion. Recognition is performed using these histograms. The diagram of the proposed system is given in Figure 1.
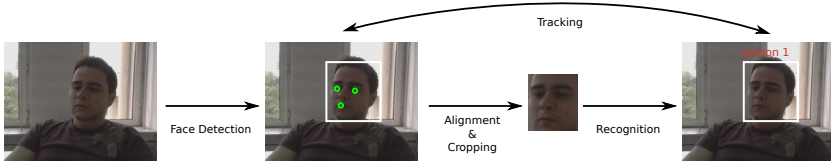

Fig. 1. Illustration of the proposed system

2 Face Detection

The face detection is performed using the features obtained through MCT transformation [1]. A strong face/non-face classifier is constructed by cascading 4 AdaBoost classifiers that use MCT-based features as weak classifiers. Multiscale face-detection is achieved by constructing a Gaussian image pyramid and searching for faces across each level of this pyramid using coarse sampling [2]. Multiview face detection is achieved by training three separate face/non-face classifiers for each pose: left profile, frontal, right profile. Each classifier is practically robust against 30° of in-plane rotation ($[-15°, +15°]$). Two additional classifiers are trained for each pose to handle in-plane rotation within the ranges of $[-45°, -15°]$ and $[+15°, +45°]$, covering 90° of in-plane rotation in total.

3 Tracking

Kalman filtering-based tracking is employed in order to link the faces in time domain. A face model is assigned for each face, and its location is updated at each frame with the location of the relevant new detection if exists. Face tracking also helps maintaining a smooth transition between frames, and eliminating many false alarms that appear on single frames.

4 Facial Feature Detection and Face Alignment

The faces must be aligned with a sufficient accuracy before they are sent to the recognition process. In order to register the faces, first, the eyes are detected within the face. Then, the faces are de-rotated so that the eyes are aligned. Next, the mouth is detected inside the de-rotated face and finally, the face is cropped using the locations of the eyes and the mouth. MCT based classifiers are used for facial feature detection, as they are used for face detection.
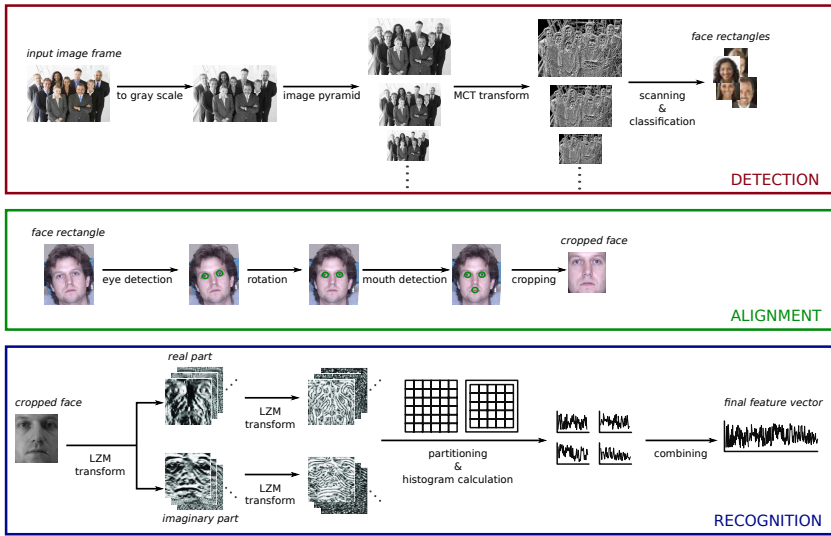
Fig. 2. Detailed diagrams corresponding to different tasks

5 Recognition

The LZM representation is used to perform face recognition. According to this technique, face images are first decomposed into a set of complex moment components using the LZM transformation. Once the moment components are obtained, they are partitioned into subregions and the phase-magnitude histograms are computed at each subregion of each component. The final feature vectors are obtained by simply concatenating all of the local phase-magnitude histograms. Once the feature vectors of the faces are obtained, they are compared using L_1 norm.
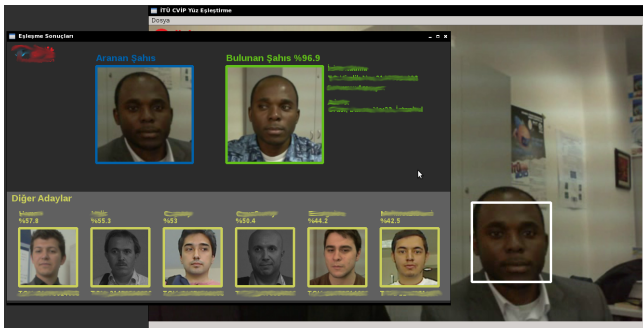

Fig. 3. A screenshot from face matching mode. The detailed matching results are displayed when the interested face rectangle is selected.
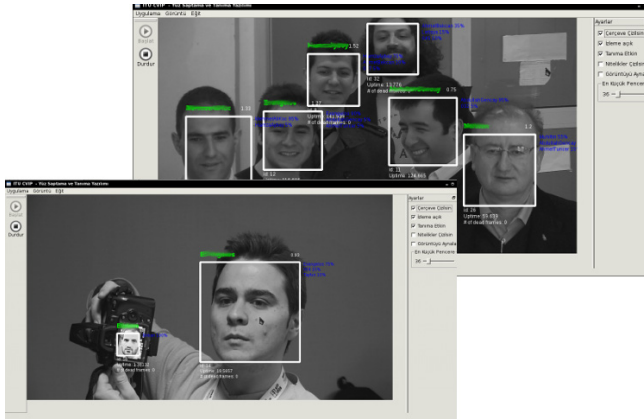
Fig. 4. Screenshots from the realtime mode. The detailed matching results are displayed next to each face rectangle.

6 Real-Life Results

The system is developed as a desktop application which can be run on regular PCs. Two different modes are available to perform recognition on still images and videos. Figure 3 and Figure 4 demonstrate example usages.

References

1. Fröba, B., Ernst, A.: Face detection with the modified census transform. In: Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 91–96. IEEE Computer Society, Washington, DC (2004)
2. Sarıyanidi, E., Tek, S.C., Gökmen, M.: Efficient face detection using coarse sampling. In: Conference on Signal Processing and Communications Applications (SIU), Turkey (2011)
3. Sarıyanidi, E., Dağlı, V., Tek, S.C., Tunç, B., Gökmen, M.: Local zernike moments: A new representation for face recognition. In: International Conference on Image Processing, ICIP (accepted, 2012)