

Face Detection and Tracking in Video Sequence using Fuzzy Geometric Face Model and Motion Estimation

P. S. Hiremath

Department of Computer Science
Gulbarga University, Gulbarga-585106
Karnataka, India

Manjunath Hiremath

Department of Computer Science
Gulbarga University, Gulbarga-585106
Karnataka, India

Mahesh R.

Department of Computer Science
Gulbarga University, Gulbarga-585106
Karnataka, India
maheshswamy99@gmail.com

ABSTRACT

With advances in computing and telecommunications technologies, digital images and video are playing key roles in the present information era. Human face is an important biometric object in image and video databases of surveillance systems. Detecting and locating human faces and facial features in an image or image sequence are important tasks in dynamic environments, such as videos, where noise conditions, illuminations, locations of subjects and pose can vary significantly from frame to frame. In this paper, a novel approach of the detection and tracking of face in video sequence based on the fuzzy geometrical face model and motion estimation is presented. The feature extraction process is performed in the support region which is determined by the fuzzy rules to detect face in an image frame. Then, the consecutive frames from a video and their corresponding optical flow are estimated, which are used for tracking face in the video sequence. The experimental results demonstrate the efficacy of the proposed method.

Keywords:

Face detection, Fuzzy geometric face model, Motion estimation, Tracking.ifx

1. INTRODUCTION

With advances in computing and telecommunications technologies, digital images and video are playing key roles in the present information era. The huge amount of visual information is handled by image and video databases, which require effective and efficient mechanisms to index and search these imagery data. In recent years, techniques have been proposed allowing users to search images by visual features, such as texture, color, shape, and sketch, besides traditional textual keywords.

Human face is an important biometric object to be searched in image and video databases of surveillance systems. Since face is a unique feature of human beings, and is ubiquitous in photos, news video, documentaries, etc., faces can be used to index and search the image/video databases, classify video scenes, and segment human objects from the background. Therefore, research on face detection has far reaching consequences in image and video database applications.

In general image and video databases, however, there is little or no constraint on the number, location, size, and orientation of human faces in the scenes. The backgrounds of these images and video sequences are usually cluttered. Thus, successful face detection and tracking becomes important and challenging before the indexing, search, and recognition of the faces could be done. Detecting and locating human faces and facial features in an image or image sequences are important tasks in dynamic environments, such as videos, where noise conditions, illumina-

tions, locations of subjects and pose can vary significantly from frame to frame.

A survey of literature on the research work focusing on various potential problems and challenges in the face detection and tracking can be found in [1,2]. There have been various approaches proposed for face detection, which could be generally classified into four categories[3]: template matching based methods, feature-based methods; knowledge-based methods, and learning based methods. Template matching based method means the final decision comes from the similarity measurement between input image and the template. It is scale-dependent, rotation-dependent and computationally expensive. Feature-based methods use low-level features such as intensity[4], color[5], edge, shape[6], and texture to locate facial features, and further, find out the face locations. Knowledge-based methods[7] detected an isosceles triangle (for frontal view) or a right triangle (for side view). Learning based methods use a lot of training samples to make the classifier to be capable of judging face from non-face. Despite of the notable successes achieved in the past decades, making a tradeoff between computational complexity and detection efficiency is the main challenge. Among the face detection algorithms, skin color based detection information is an important category[8]. Hiremath and Manjunath[9] have proposed fuzzy geometric approach for face model construction based on only two features, namely, eyes and mouth, for still images which are shown to be the optimal discriminating features for face detection.

In this paper, a novel method is proposed for face detection and tracking in a video sequence by using fuzzy geometric face model and motion estimation. The effectiveness of the proposed method is demonstrated by the experimental results. The experimentation has been done using publicly available video database.

2. MATERIALS AND METHODS

The Honda/UCSD Video Database provides a standard video database for evaluating face detection, tracking and recognition algorithms. Each video sequence is recorded in an indoor environment at 15 frames per second, and each sequence lasted for at least 15 seconds. The resolution of each video sequence is 640x480. Every individual is recorded in at least two video sequences. In each video, the person rotates and turns his/her head in his/her own preferred order and speed, and typically in about 15 seconds, the individual is able to provide a wide range of different poses.

The Honda/UCSD Video Database contains two datasets. The first dataset is recorded by a SONY EVI-D30 camera at Honda Research Institute in 2002. It includes three different subsets, one each for training, testing, and occlusion testing. Each subset

contains 20, 42, 13 videos respectively from 20 human subjects. The second dataset is recorded by a SONY DFW-V500 camera at Computer Vision Laboratory, University of California, San Diego, in 2004. It includes two subsets, one each for training and testing, of 30 videos from another 15 different human subjects [10].

3. PROPOSED METHODOLOGY

The proposed methodology comprises the application of fuzzy geometric face model for face detection and motion estimation for face tracking in video sequences, which are described below.

3.1 Fuzzy Geometric Face Model for Detection

Firstly, different frames of the input video sequence are extracted. Then, fuzzy geometric face model for face detection[9] is applied. The extracted frame image is preprocessed and then the eyes are searched on the basis of geometrical knowledge of the symmetrical relations between eyes. The other prominent feature, namely, mouth, is searched with respect to the detected eyes using fuzzy rules and the face detection algorithm. The fuzzy rules are derived from the knowledge of the relative positions of the facial features in the human faces and the trapezoidal fuzzy membership functions to represent the uncertainty of the locations of the facial features due to variations in poses and facial expressions. The frames of the input video sequence are expected to contain not too dark or too bright. If the input frame of the video sequence is a color frame, it is converted into gray scale image. The gray scale frame is filtered using the Sobel horizontal edge emphasizing filter and utilizing the smoothing effect by approximating a vertical gradient. In the filtered frame, objects of interest (facial features) are brighter than the background. In order to make the essential facial features clearly visible, filtered image is converted into binary frame by simple global thresholding. Further, the frame is denoised by morphological operations, in which opening operation is performed to remove noise, and then the closing operation is performed to remove holes. Then the active pixels are grouped into maximal connected blocks to get the regions or blocks which are labeled. After the labeling process, for each feature block, its center of mass(x, y), orientation θ , bounding rectangle and the length of semi major axis are computed[9]. The output of the face detection algorithm using fuzzy geometric face model is shown in the Fig. 1.

3.2 Motion Estimation for Face Tracking

Motion estimation is the process of determining motion vectors that describe the transformation from one 2D image to another; usually from adjacent frames in a video sequence. The motion vectors may relate to the whole image (global motion estimation) or specific parts, such as rectangular blocks, arbitrary shaped patches or even per pixel. The motion vectors may be represented by a translational model or many other models that can approximate the motion of a real video camera, such as rotation and translation in all three dimensions and zoom.

Optical flow reflects the image changes due to motion during a time interval dt , and the optical flow field is the velocity field that represents the three-dimensional motion of the object points across a two-dimensional frame. Optical flow is an abstraction used by computational methods for motion estimation. Therefore, it should represent only those motion-related intensity changes in the frame that are required in further processing, and all other image changes in the optical flow should be considered errors of flow detection[11-13].

Assuming a continuous frame, let $f(x, y, t)$ refer to the gray-level of (x, y) at time t representing a dynamic frame as a function of

position and time, which permits it to be expressed as a Taylor series:

$$f(x+dx, y+dy, t+dt) = f(x, y, t) + f_x dx + f_y dy + f_t dt + O(\partial^2) \quad (1)$$

where f_x, f_y, f_z denote the partial derivatives of f with respect to x, y, t respectively. One can assume that the immediate neighborhood of (x, y) is translated some small distance (dx, dy) during the interval dt ; that is, dx, dy, dt can be found, such that

$$f(x + dx, y + dy, t + dt) = f(x, y, t) \quad (2)$$

If dx, dy, dt are very small, the higher-order terms in the equation (1) vanish and

$$-f_t = f_x \frac{dx}{dt} + f_y \frac{dy}{dt} \quad (3)$$

The goal is to compute the velocity

$$c = \left(\frac{dx}{dt}, \frac{dy}{dt} \right) = (u, v) \quad (4)$$

The quantities f_x, f_y, f_t can be computed, or at least approximated, from $f(x, y, t)$. The motion velocity c can then be estimated as

$$-f_t = f_x u + f_y v = \text{grad}(f)c \quad (5)$$

where $\text{grad}(f)$ is a two-dimensional frame image gradient. The Fig.2 shows the intermediate motion estimation between frames.

3.3 Proposed Method

The block diagram of the proposed method for face detection and tracking in a video sequence is shown in the Fig. 3

The algorithm of the proposed method is given below:
Algorithm : Face detection and tracking

- (1) Input video sequence.
- (2) Extract all the frames from the input video sequence, and then select first video frame as key frame.
- (3) Apply the fuzzy geometric face model[9] for searching the face region in the key frame using prominent face features, namely, eyes and mouth, to detect the face.
- (4) Select the next video frame and perform the motion estimation to determine the direction and distance of object motion from one video frame to the next frame.
- (5) Draw the rectangular box for the detected face in the frame.
- (6) Repeat the Step 4 and 5 till the end of the input video sequence, which results in tracking the detected face in the video sequence.

4. EXPERIMENTAL RESULTS AND DISCUSSION

The experimentation of the proposed approach is carried out using the Honda/UCSD Video database[10]. The implementation is done on Intel Core2Quad PC @ 2.60 GHz machine using MATLAB 7.0. The different frames of the input video sequence are extracted. The extracted frame image is preprocessed and then the facial features, namely, eyes and mouth are searched by using fuzzy geometric face model[9], and then, the detected face is tracked by using motion estimation[11-13]. The experimental results of the face detection and tracking in video sequence by the proposed method is shown in the Fig. 4

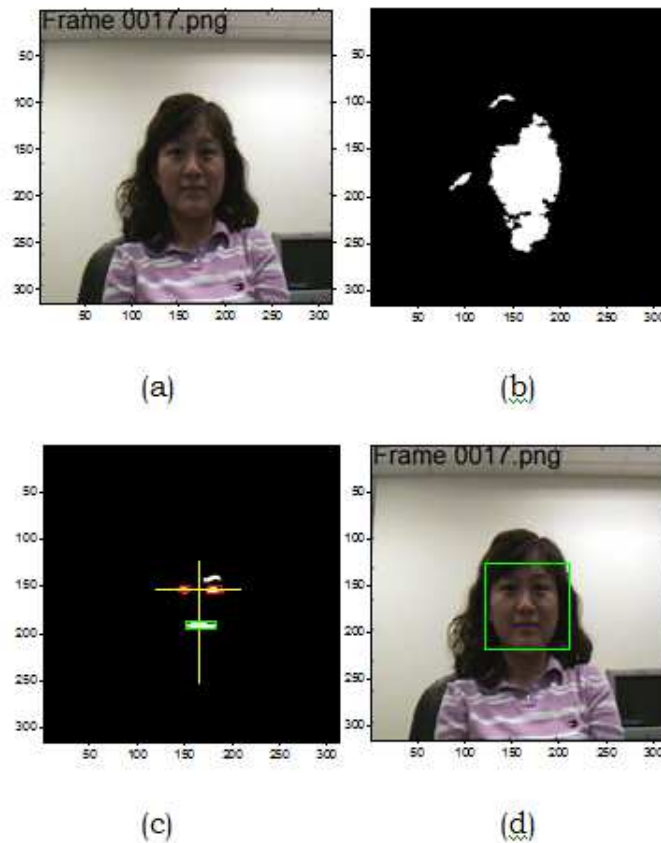


Fig. 1. Face detection using fuzzy geometric face model (a) Original input image (b) Skin region extraction (c) Construction of fuzzy face model (d) detected face



Fig. 2. Two consecutive frames from a video and their corresponding optical flow quiver plot

The comparison of the tracking results obtained by the proposed method and other methods are shown in Table 1.

Table 1. Comparison of the tracking results obtained by the proposed method and other methods

Parameters	Shaohua Zhou et al. [14]	Proposed Method
Video	Face	Face
#video frames	800	395
Frame rate	-	15 fps
Frame Size	240x360	640x480
Occlusion	Yes	Yes

5. CONCLUSION

In this paper, a novel method of the detection and tracking of face in video sequence based on the fuzzy geometrical face model and motion estimation is presented. The human face is detected by feature extraction process based on fuzzy geometric face model. Then, the consecutive frames from a video and their corresponding optical flow is estimated and face is tracked. In the proposed method, single frontal face in the video frames with different motions, head tilts, lighting conditions, expressions and backgrounds are considered. The proposed approach yields better average detection and tracking, which is robust and almost real time. The proposed method can be extended for multiple faces in

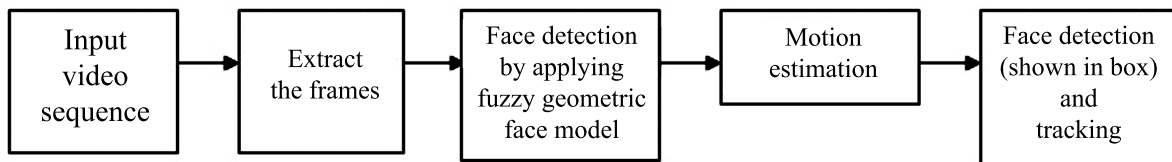


Fig. 3. Block diagram of the proposed approach

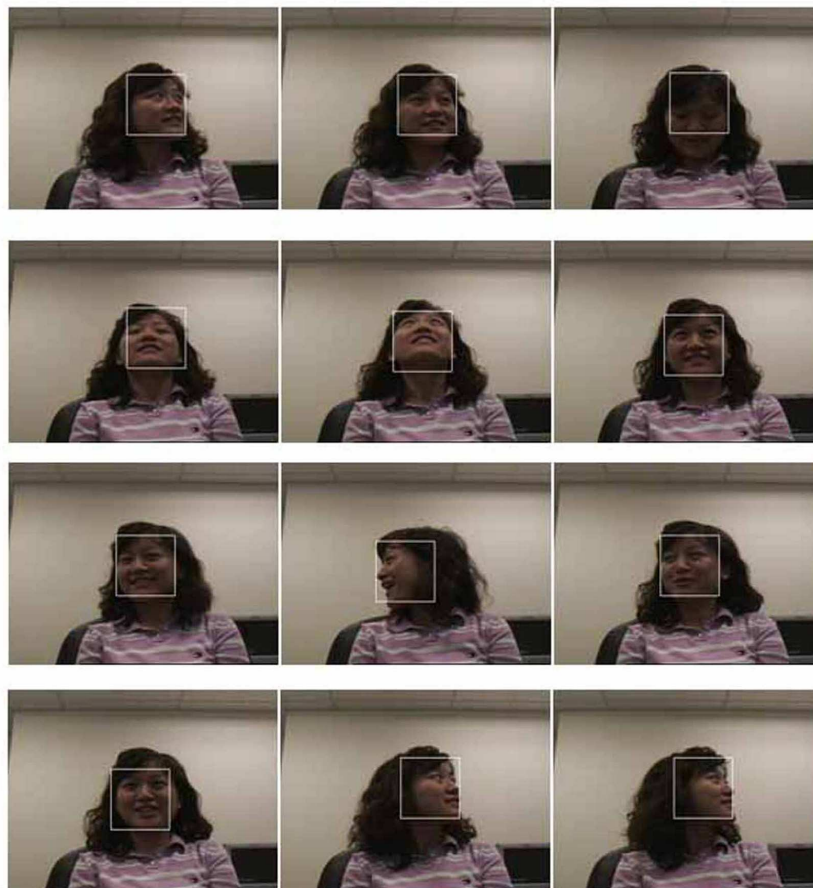


Fig. 4. The experimental results of the face detection and tracking in a sample video sequence by the proposed method

video sequence by considering multiple face detection and tracking algorithms.

6. ACKNOWLEDGEMENT

The authors are grateful to the referees for their helpful comments and suggestions. Also, the authors are indebted to the University Grants Commission, New Delhi, for the financial support for this research work under UGC-MRP F.No.39-124/2010 (SR).

7. REFERENCES

- (1) Kinjal A Joshi, Darshak G, Thakore, "A survey on moving object detection and tracking in video surveillance system", *International Journal of Soft Computing and Engineering (IJECSE)* Vol. 2, No. 3, (July 2012), pp. 44-48.
- (2) Huafeng Wang, Yunhong Wang and Yuan Cao, "Video based face recognition : A Survey", *World Academy of science, Engineering and Technology*, pp.293-301.
- (3) Yunawen Wu, Xueyi Ai, "Face Detection in Color Images Using AdaBoost Algorithm Based on Skin Color Information", *First International Workshop on Knowledge Discovery and Data Mining*, 2008, pp.339-342.
- (4) Kwok-Wai Wong, Kin-Man Lam, and Wan-Chi Siu, "An Efficient Algorithm for Human Face Detection and Facial Feature Extraction under Different Conditions", *Pattern Recognition*, Vol. 34, 2001, pp.1993-2005.
- (5) Ing-Sheen Hsieh, Kuo-Chin Fan, and Chiunhsiun Lin, "A Statistic Approach to the Detection of Human Faces in Color Nature Scene", *Pattern Recognition*, 35, 2002, pp.1583-1596.
- (6) J. G. Wang and T. N. Tan, "A New Face Detection Method Based on Shape Information", *IEEE Transactions on Pattern Recognition Letters*, Vol. 21, 2000, pp. 463-471.
- (7) Chiunhsiun Lin and Kuo-Chin Fan, "Triangle-based Approach to the Detection of Human face", *Pattern Recognition*, Vol. 34, No. 6, 2001, pp.1271-1283.
- (8) Hsu, R. L., Mottaleb, M. A., Jain, A. K. "Face detection in color images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 5, 2002, pp.696-706.

- (9) P. S. Hiremath and Manjunath Hiremath, "Fuzzy face model for face detection using eyes and mouth features", *International Journal of Machine Intelligence*, Vol. 3, No. 4, 2011, pp.185-190.
- (10) Kuang-Chih Lee, Jeffrey Ho, Ming-Hsuan Yang, David Kriegman, "Visual tracking and recognition using probabilistic appearance manifolds", *Computer Vision and Image Understanding*, Vol. 99, 2005, pp.303–331.
- (11) Milan Sonka , Vaclav Hlavac , Roger Boyle, "Image Processing: Analysis and Machine Vision", Second Edition, Thompson Learning, 2008.
- (12) Stan Z. Li, Anil K. Jain, "Handbook of face recognition", Second Edition, Springer-verlag, 2011, ISBN:978-0-85729-931-4.
- (13) Philip H.S. Torr and Andrew Zisserman, "Feature Based Methods for Structure and Motion Estimation", *ICCV Workshop on Vision Algorithms*, 1999, pp.278-294.
- (14) Shaohua Zhou, Rama Chellappa, Baback Moghaddam, "Visual Tracking and Recognition Using Appearance-Adaptive Models in Particle Filters", *IEEE Transactions on Image Processing*, Vol. 13, No. 11, 2004, pp. 1491-1506.