Comprehensive Survey On Face Tracking In Real Time Video

Amit Bhoomkar, Ajay Kulkarni and Somnath Gosavi .
Department of Computer Engineering, V.I.I.T., Pune University
amitbhoomkar@gmail.com ,ajay31514@gmail.com,gosavi.somesh@gmail.com

Abstract:-Automatic face identification of suspects in real time video has drawn significant research interests and led to many interesting applications. It is a challenging problem due to the huge variation in the appearance of each person in the video. Although existing methods demonstrate promising results in clean environment, the performances are limited in complex videos due to the noises generated during the face tracking and face clustering process. This paper provides an up-to-date critical survey of still- and video-based face recognition researches as well as different methods of face recognition and their pros and cons.

Special Keywords—Face Recognition, Facial Tracking, Face Segmentation, Face Modeling, Face Verification, Feature extraction.

I. INTRODUCTION

Image processing or widely called digital image processing were developed in 1960s and among the earliest developers are Jet propulsion laboratory, MIT and Bell Labs. It was originally developed for satellite imagery, medical imaging, videophone, photo enhancement and character recognition. At the beginning, the cost of computing was too high so it was only until 1970s that the area when the cheaper computers are available with dedicated hardware.

Now, in present era, computer with developed technologies and signal processors of digital image processing has become the most common form of image processing due to its versatility and low cost. In simple words, image processing is the ability to extract information from images. To extract information of any kind from an image, image has to be transformed into a data set and then mathematical operations can be applied to it. There are numerous programs that handle image processing. MATLAB is one of the most popular programs for image processing.

Image processing involves intensive mathematics and MATLAB is designed for intensive mathematics, therefore it is ideal program for writing image processing applications. Unlike MATLAB, IMAGEMAGIC is another program that can do image processing.

II. IMAGE TRACKING

For image tracking application of image processing requires face detection and recognition. When the faces are located by face recognition technique, their features can be tracked. Face tracking is very critical process of image processing. Tracking plays a key role in spatiotemporal-based recognition methods.

In general terms, tracking is essentially motion estimation. Tracking is constrained by a generic 3D model or a learned statistical model under deformation. Face tracking roughly divided into three categories:

(1) Head Tracking- It involves tracking the motion of rigid object that is performing rotations and translations.

(2) Facial Tracking- It involves tracking of non-rigid deformations. It also involves articulated motion due to speech or facial expressions and deformable motion due to muscle contractions and relaxations.

(3) Complete Tracking- It involves tracking in both the heads and facial features.

Most of the efforts had been taken on first two problems: head tracking and facial feature tracking. Head tracking is generally carried out with the high Hessian values. Several such values cause points on head which are tracked and 3D motion parameters of head are reconverted by solving over constrained set of motion equations.

Facial feature tracking method may involve the use of feature boundary or the feature region. Feature boundary deals with the accuracy of the facial shape. For example, to track the boundaries of lips and mouth this method is used.

A face tracking system was used in Maurer and Malsburg [1996b]. This system used a graph representation about 20-40 nodes to model the face. Two tracking system described in Jebara et al [1998] and Strom et al [1999] which involves model faces completely filled with texture and geometry. Some of the newer model of tracking can calculate the 3D motions and deformations directly from image intensities.

III. FACE RECOGNITION

- ANALYSIS OF HUMAN FACE

Psychologically a face is rich source of information about human behavior. It is the feature which distinguishes a person. Geometrically a face is 3D spaces sum of large number of polygons that can be represent by pixel.

- FACE RECOGNITION BY HUMAN

How do we recognize familiar faces from a variety of viewpoints? This is a fundamental problem from both psychological and computational perspectives. With the help of brain human can recognize the face. Many of the neurologists believe that the brain identifies faces as "special" and very different from the other visual objects. The human brain combines motion and shape information to recognize faces and facial expressions. More recent studies have suggested that there may be particular neurons tuned to the identity of one particular person.

- FACE RECOGNITION BY COMPUTER

Face recognition has recently received significant attention, especially during past few years. There are at least two reasons for this trend; first is the wide range of commercial and law enforcement applications and second is availability of feasible technologies.

Face recognition techniques roughly classify into two groups depending on whether they make use of static images or video. A general statement of the problem of face recognition through computer can be formulated as follows: given a still or video image of a scene, identify one or more persons in the scene using a stored database faces.

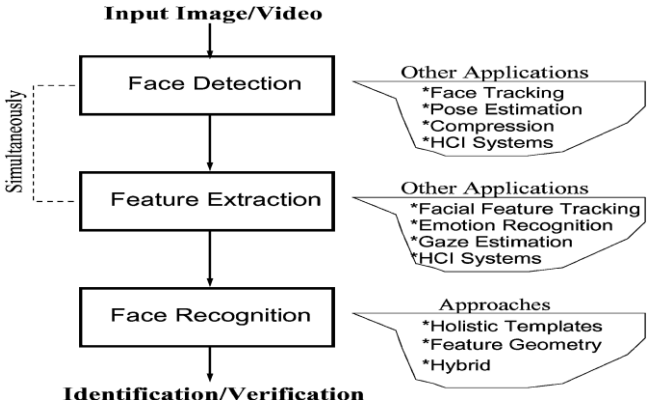

Fig. 1 Configuration of a face recognition system.

[8]

- FACE RECOGNITION FROM STILL IMAGES

Automatic face recognition involves three key steps:
(1) Face detection
(2) Feature extraction
(3) Face identification or verification.

Sometimes face detection and feature extraction can be achieved simultaneously depending on the nature of application. Automatic face recognition system must perform all the three tasks. Earlier face detection techniques could only handle single or a few well-separated frontal faces in images with simple backgrounds. Extensive research on each tasks has been carried out because of which face detection can possible at complex stages also.

Key steps to perform automatic face detection- 1) the detection of faces in images. After a face has been detected, task of feature extraction is carried out. Depending on the type of classification, system features can be local features such as eyes, nose and mouth. Feature extraction is also required for animation and recognition of facial expressions. As compare to invariant-feature –based methods, multi-view-based methods of face detection seem to be able to achieve better results.

As suggested by studies in psychology, face recognition systems need facial features in addition to the holistic face. There are three types of feature extraction methods can be distinguished:
(1) Generic methods based on edges, lines and curves;
(2) Feature template-based methods that are used to detect facial features such as eyes;
(3) Structural matching methods that take into consideration geometrical constraints on the features.

Early days, approaches focused on individuals, but now combination of all types comes into picture. Combination of various tasks produces challenge for the feature extraction. An even more challenging situation for feature extraction is feature 'restoration', which tries to recover features that are invisible due to large variations in head pose.

Face recognition based on still images can be viewed as 2D image matching and recognition. This application is not beneficial in most commercial/law enforcement fields. Significant program has been achieved on various aspects of face recognition: segmentation, feature extraction and identification of faces. More program also been made on constructing fully automatic systems that integrate all these techniques.

Though machine recognition of faces from still images has achieved a certain level of success, its performance is still far from that of human perception. Hybrid face recognition system which uses both holistic and local features resembles the human perceptual system. However, many questions need to be answered before we can build such a combined system. One of the important question is how to arbitrate the use of holistic and local features? Still the research is going on to find the easiest way to use the combined system. The challenge of developing face detection technique that report not only the presence of face but also the accurate locations of facial features, Without accurate localization of important features, accurate and robust face recognition cannot be achieved.

- FACE RECOGNITION FROM VIDEO

A typical video-based face recognition system automatically detects face regions, extracts features from the video, and recognizes facial identity if a face is present. In surveillance, information security, and access control applications, face recognition and identification from a video sequence is an important problem. Significant challenges for video-based recognition still exist such as low quality of video, small images of face etc.

- TECHNIQUES OF VIDEO BASED RECOGNITION

Roughly four computer vision areas are important for video based face recognition: segmentation of moving objects (humans) from a video sequences structure extensions, 3D models for faces and non-rigid motion analysis. As this research goes on, finally result suggests that three techniques are mainly required instead of four. These three techniques are
1) Face Segmentation and pose estimation
2) Face Tracking
3) Face Modeling

Face segmentation & pose estimation involves segmentation of moving objects (faces) from the image sequence. It uses the method of simple pixel-based change detection procedures based on different images. More recent methods have used motion or color information to speed up the process of searching for possible face. Stile-image-based face detection techniques can be applied after the candidate face regions are located. The location of feature points can be used for pose estimation, which is important for synthesizing a frontal view. Latest segmentation models locate the face and estimate it's pose simultaneously without extracting features.

Face tracking method involves the tracking of faces or complete object from the video. Already the information related to tracking is stated section 2. Now a days tracking system is also used to recognize facial expressions. This systems are based on local parameterized models.

Face modeling involves modeling of 3D shape faces and texture modeling. 3D models of faces have been employed in the graphics, animation and model based image compressions literature. Complicated face modeling systems are used in forensic face reconstruction from partial information. In Blanz and Vertter [1999], real rime 3D modeling and tracking of faces was described. A 3D model was aligned to match frontal views of face in a video. Recently multi view based 2D methods are very commonly use.

Video face recognition originated from still-image based techniques. This system automatically detects and segments the face from the video, and then applies still-image face recognition techniques. Tracking is the improvement to those systems. Use of multimodal cues is the next phase of video-based face recognition. Important of multimodal is that it uses multimodal cues which offers a comprehensive solution to the task of identification that might not be achievable by using face images alone.

More recently a third phase of video face recognition has started. These methods coherently exploit both spatial information and temporal information. In Wechsler et al [1997], a fully automatic person authentication system was described. Which include both video break, face detection and authentication modules. This system was tested on three image sequences: the first was taken indoors with the subject presents, the second was taken outdoors with the two subjects, and the third was taken outdoors with one subject under stormy conditions. Perfect results were reported on all three sequences, as verified against a database of 20 still face images.

An appearance model based video tracking and enhancing identification was proposed by Edwards et al [1998]. Appearance model is the combination of active shape model and the shape-free texture model after cutting the face into a mean shape.

The key to building a successful video based system is use temporal information to compensate for the lost spatial information. There are various issues involves in face detection from video such as quality of video, limited database good hardware etc. The databases used in many systems have been small. This is happened due to the tremendous amount of storage space needed for video sequences. However during the past 8 years a lot of research had been completed to overcome these issues. Recognition of human behavior has been actively studied.

Typical Applications of Face Recognition

Areas	Specific applications
Entertainment	Video game, virtual reality, training programs Human-robot-interaction, human-computer-interaction
Smart cards	Driver's license, entitlement program immigration, national ID, passports, voter registration welfare fraud
Information security	TV Parental control, personal device logon, desktop logon, Information security, Application security, database security, file encryption, Intranet security, internet access, medical records Secure trading terminals

Law enforcement	Advanced video surveillance, CCTV control and surveillance Portal control, post event analysis Shoplifting, suspect tracking and investigation

[8]

Table II. Available Commercial Face Recognition Systems (Some of these Web sites may have changed or been removed.)The identification of any company, commercial product, or trade name does not imply endorsement or recommendation by the National Institute of Standards and Technology or any of the authors or their institutions.

Commercial products	**Websites**
FaceIt from Visionics	http://www.FaceIt.com
Viisage Technology	http://www.viisage.com
FaceVACS from Plettac	http://www.plettac-electronics.com
FaceKey Corp.	http://www.facekey.com
Cognitec Systems	http://www.cognitec-systems.de
Keyware Technologies	http://www.keywareusa.com/
Passfaces from ID-arts	http://www.id-arts.com/
ImageWare Sofware	http://www.iwsinc.com/
Eyematic Interfaces Inc	http://www.eyematic.com/
BioID sensor fusion	http://www.bioid.com

[8]

III. CONLCUSION

To conclude our paper, we present opinion about face recognition based on psychological studies and lessons learned from designing algorithms. We conclude that different mechanisms are involved in human recognition of familiar and unfamiliar faces. In this paper we have presented an extensive survey of machine recognition of human faces and a brief review of related psychological studies. We have considered two types of face recognition tasks: one from still images and the other from video. We have categorized the methods used for each type, and discussed their characteristics and their pros and cons. In addition to a detailed review of representative work, we have provided summaries of current developments and of challenging issues.

[8]

IV. REFERENCES

(1) Bailly-Bailliere, E., Bengio, S., Bimbot, F., Hamouz, M., Kittler, J., Mariethoz, J., Matas, J.,Messer, K., Popovici,V., Poree, F., Ruiz, B., And Thiran, J. P. 2003. The Banca Database And Evaluation Protocol. In *Proceedings Of The International Conference On Audio- And Video-Based Biometric Person Authentication*. 625–638.

(2) Cox, I. J., Ghosn, J., And Yianilos, P. N. 1996. Feature-Based Face Recognition Using Mixture Distance. In *Proceedings, Ieee Conference On Computer Vision And Pattern Recognition*. 209–216.

(3) Jebara, T., Russel, K., And Pentland, A. 1998. Mixture Of Eigen Features For Real-Time Structure From Texture. Tech. Rep. Tr-440, Mit Media Lab, Massachusetts Institute Of Technology, Cambridge, Ma.

(4) Hancock, P., Bruce, V., And Burton, M. 1998. A Comparison Of Two Computer-Based Face Recognition Systems With Human Perceptions Of Faces.*Vis. Res. 38*, 2277–2288.

(5) Lanitis, A., Taylor, C. J., And Cootes, T. F. 1995. Automatic Face Identification System Using Flexible Appearance Models. *Image Vis. Comput. 13*,393–401.

(6) Schneiderman, H. Andkanade, T. 2000. Probabilistic Modelling Of Local Appearance And Spatial Relationships For Object Recognition. In *Proceedings. Ieee Conference On Computer Vision And Pattern Recognition*. 746–751.

(7) Phillips, P. J., Wechsler, H., Huang, J., And Rauss, P. 1998b. The Feret Database And Evaluation Procedure For Face-Recognition Algorithms. *Image Vis. Comput. 16*, 295–306.

(8) Face Recognition: A Literature Survey By W. Zhao R. Chellappa, P. J. Phillips, A. Rosenfeld

(9) Y. Zhang, C. Xu, H. Lu, And Y. Huang, "Character Identification In Feature-Length Films Using Global Face-Name Matching," *Ieee Trans. Multimedia*, Vol. 11, No. 7, Pp. 1276–1288, November 2009.

(10) J. Stallkamp, H. K. Ekenel, And R. Stiefelhagen, "Video-Based Face Recognition On Real-World Data." In *Iccv*, 2007, Pp. 1–8

(11) Gong, S., Mckenna, S., And Psarrou, A. 2000. *Dynamic Vision: From Images To Face Recognition.* World Scientific, Singapore.

(12) Li, B. And Chellappa, R. 2001. Face Verification Through Tracking Facial Features. *J. Opt. Soc. Am.18.*

(13) Swets, D. L. And Weng, J. 1996. Discriminant Analysis And Eigen Space Partition Tree For Face And Object Recognition From Views. In *Proceedings, International Conference On Automatic Face And Gesture Recognition*. 192–197.