Available online at www.elixirpublishers.com (Elixir International Journal)

Computer Science and Engineering

ISSN: 2229-712X

Elixir Comp. Sci. & Engg. 58 (2013) 15008-15011

Automatic facial detection from input videos

S.Muralidharan[1] and T. Dhavaselvan[2]

[1]Department of Computer Science, Dravidian University, Kuppam, A.P, India.

[2]Department of Computer Science, Pondicherry University, Pondicherry, India.

ARTICLE INFO

Article history:
Received: 7 March 2013;
Received in revised form:
15 May 2013;
Accepted: 17 May 2013;

Keywords
Face detection,
Facial feature extraction,
Face identification.

ABSTRACT

Recognizing faces is something that people usually do effortlessly and without much conscious thought, yet it has remained a difficult problem in the area of computer vision, where some 20 years of research is just beginning to yield useful technological solutions. As a biometric technology, automated face recognition has a number of desirable properties that are driving research into practical techniques. The problem of face recognition can be stated as 'identifying an individual from images of the face' and encompasses a number of variations other than the most familiar Application of mug shot identification. One notable aspect of face recognition is the broad interdisciplinary nature of the interest in it: within computer recognition and pattern recognition; biometrics and security; multimedia processing; psychology and neuroscience. It is a field of research notable for the necessity and the richness of interaction between computer scientists and psychologists. The automatic recognition of human faces spans a variety of different technologies. At a highest level, the technologies are best distinguished by the input medium that is used, whether visible light, infra-red or 3-dimensional data from stereo or other range-finding technologies. Thus far, the field has concentrated on still, visible-light, photographic images, often black and white, though much interest is now beginning to be shown in the recognition of faces in colour video. Each input medium that is used for face recognition brings robustness to certain conditions, e.g. infra-red face imaging is practically invariant to lighting conditions while 3-dimensional data in theory is invariant to head pose. Imaging in the visible light spectrum, however, will remain the preeminent domain for research and application of face recognition because of the vast quantity of legacy data and the ubiquity and cheapness of photographic capture equipment.

1. Introduction

Face detection is the first processing stage in a face-recognition system. It refers to the process which segments possible faces from their background. This process takes an image (or a video sequence) as input, and returns the *location* and *scale* of each face, if there exists any. This is of crucial importance for unconstrained face recognition, where no *a-prior* knowledge is available as to the locations and scales of possible faces in an image. In this Paper, we consider detectors that are capable of finding multiple faces in one image. In addition to face recognition, face detection is also an important research topic for a number of reasons.

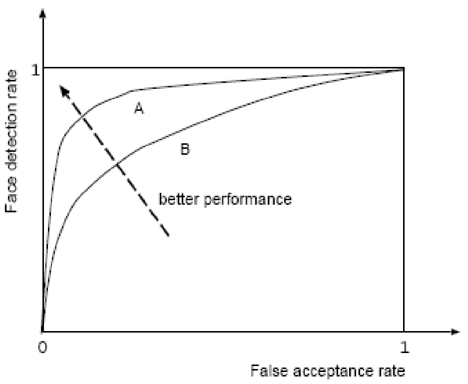

Figure 1: Face detection accuracy curve

1.1 Key elements of face-detection techniques

Tremendous effort has been spent on designing face detectors in the past decade. Various techniques have been proposed in literature, which range from simple heuristics-based algorithms to advanced machine-learning-based algorithms.

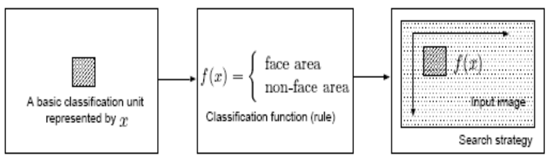

Figure 1.1: Elements in a Face Detector

1.2 Colour-based detectors

In colour-based detectors, an image pixel is regarded as the basic classification unit, which is represented by its associated colour components in certain colour spaces. Commonly-used colour spaces are normalized *RGB*, *Y CbCr* and *HSV* [56]. A typical classification function used for colour-based detectors is based on the Bayes decision rule. More specifically, assume that vector x contains the colour components of a pixel, x is classified as a skin pixel if

$$\frac{p(x\,|\,skin)}{p(x\,|\,non-skin)} > \beta \qquad \text{Equation - 1.1}$$

where $p(x/\text{skin})$ and $p(x/\text{non-skin})$ are class-conditional probability density functions and β is a constant.

Tele:
E-mail addresses: smd2008@rocketmail.com

The class-conditional probability densities can be estimated by various techniques, such as histograms, parametric modelling based on uni-modal Gaussian or mixture of Gaussians [56]. Suppose a uni-modal Gaussian is used to characterize the skin-colour distribution, the class-conditional pdf becomes

$$p(x\,|\,skin) = (2\pi)^{-d/2}\,|C|^{-1/2}\,e^{-1/2(x-m)^T}C^{-1}(x-m) \qquad \text{Equation-1.2}$$

where C denotes the covariance matrix, m is the mean colour vector and d is the dimension of the colour feature vector. Furthermore, it is assumed that the non-skin colour distribution $p(x/\text{non-skin})$ is uniform. Using this assumption formula can be derived from equation 1.1

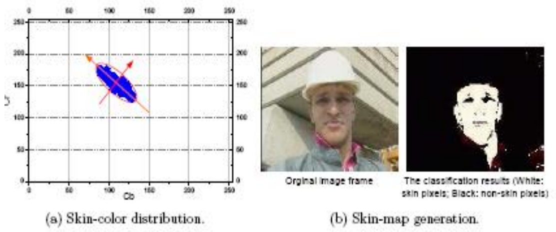

Figure 1.2 Skin colour classification using uni-modal Gaussian

$$(x-m)^T C^{-1}(x-m) < \alpha \qquad \text{Equation - 1.3}$$

Where α is a constant. This formula actually represents an ellipse in 2D or an ellipsoid in 3D. In Fig. 1.2(a), we show an elliptic modelling of the skin colour distribution in the *Cb -Cr* space. Since *C* and *m* can be estimated from a set of skin-colour samples (see Fig. 1.2(a)), and α can be empirically chosen, Eqn. (1.3) can be used as an explicit function to classify a pixel to face/non-face area.

1.3 Skin-colour detector

The skin-colour sub detector that we adopted is based on a uni-modal Gaussian modelling of skin-colour distribution in the C_b - C_r colour space. By applying the skin-colour detector to an input image, we can obtain a skin map, which is further smoothened as shown in Fig. 1.3. Each connected region in the skin map is then marked as a *skin blob* and passed on to the next detector. Note that the colour-pruning criterion used in this stage is very relaxed, resulting in a higher false acceptance rate than a more intricate colour-based detector. However, for an initial detector in the complete detector cascade, we intend to put more emphasis on speed than accuracy and the overall detection accuracy is reinforced by the succeeding detectors.

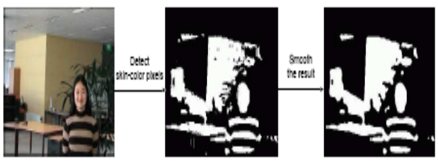

Figure 1.3 Skin- colour based face detector

2. Facial Feature Extraction

Facial feature extraction locates important feature positions within a detected face. The features which are most interesting are eyes, mouth, nose and eyebrows. In some cases, it is necessary to obtain more detailed descriptions of these features, such as their contours Accurate feature extraction facilitates *face region normalization*, where the detected face can be aligned to a common coordinate framework. This can significantly reduce the large variances introduced by different face scales and poses, which alleviates the difficulty for face identification. In addition, for some face-identification techniques, the accurate locations of feature points sampling the contours of facial features provide important input parameters for the face identification.

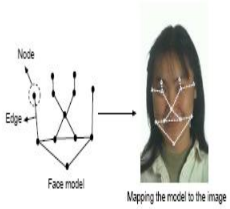

Figure 2.1 Model Based feature extraction scheme.

2.1 Model-based algorithms for locating features

A model-based algorithm is characterized by the following two important aspects.

1. ***Capture range***: A model-based algorithm usually needs an explicitly initialized model (or initial search range) to start with the model-fitting procedure.

2. ***Extraction accuracy***: For converged feature-extraction cases, The extraction accuracy is usually expressed by the deviation between a fitted model and the manually-labelled feature positions.

2.2 Deformable geometric templates

Facial features belong to a category of objects that share common shape characteristics while still having variances due to individual diversity, facial expressions, etc. A promising technique to model this kind of objects is to use deformable models, which can adapt to individual instances with certain model constraints.

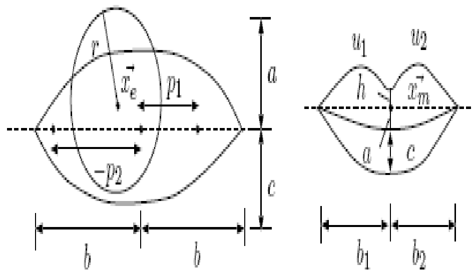

Figure 2.2 Templates for right eye and sectional view of mouth

These constraints can be based on *a-priori* knowledge of object properties.

Another technique used in literature is deformable geometric templates, which incorporate more object-specific *a-priori* knowledge about facial features. In the two parameterized feature templates (models) are pre-defined for matching the eyes and mouth individually (see Fig. 2.2). In order to find the best parameters for the model to match a given image, the technique first chooses several image representations (valley map, peak map and gradient map) to extract properties such as peaks, valleys and gradients in the gray scale values of the image. Based on these representations, an energy function *E* is then defined as a weighted sum of terms reflecting the knowledge about the feature characteristics. For instance, suppose that $V(x; y)$ is the valley map of the image, an energy term *Ev* can be defined as

$$E_v = \frac{1}{Area}\iint_{Circle-area} V(x,y)d(x,y) \qquad \text{Equation - 2.1}$$

where the integral above is performed over the interior of the eye circle.

The target position, the fitting algorithm applies an iterative procedure such as a gradient-descent algorithm to minimize the energy function *E*. It is reported in a paper that the algorithm has achieved accurate extraction results.

To improve the deformable template modelling, in a preliminary work we proposed a multistage feature-localization technique incorporating

1. A pre-processing step to select multiple initializations for feature template matching; and
2. A post processing step to verify the feature locations using a set of heuristic rules. The proposed method improves the reliability of the template matching in the existing methods.

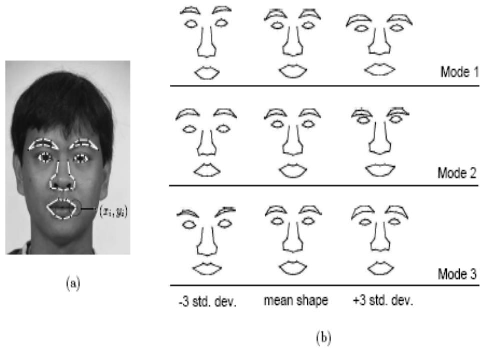

Figure 2.3: (a) Shape representation.
(b) Statistical shape deformation with the first three modes (by std. deviation).

The *a-priori* knowledge about facial features in deformable templates is specifically defined for eyes and mouth, which lacks flexibility, to deal with more general feature specifications. An alternative approach is to implicitly define facial feature properties by using statistics, which automatically derives feature models by the use of a set of training samples. The statistics-based approach for feature extraction offers more flexibility, which also can be easily applied to other object structures.

3. Face Identification

Face identification generates the final output of the complete face-recognition system: the identity of the given face image. The normalized face image and the facial feature locations derived from the previous stages, is generated from the given face and compared with a database of known faces. If a close match is found, the algorithm returns the associated identity. A key problem in face identification is the large differences between face images from the same person (*intra-personal variances*) as compared to those from different persons (*inter-personal variances*).

Face identification has a wide range of applications. Because it offers a non-intrusive way for human identification, the face is used as an important biometric in security applications. Recently, face recognition has received wide interest in border control, and a number of countries are integrating facial information into the *electronic passport* in addition to several other biometrics such as fingerprints and iris .

In order to achieve a performance face-recognition system, each processing stage in the system has to be carefully designed to satisfy specific application requirements.

4. Performance evaluation of active shape model for facial feature extraction

To demonstrate the performance gain of H-ASM, we use a training set of 200 face images to derive the shape model and a separate test set containing another 200 face images to test the extraction performance.

For each test image, we randomly perturb the centre position of the initial shape from its ground-truth position by up to 50 pixels. We perform the model fitting using both ASM and H-ASM. The results are depicted in Fig. 4.1 At the left of Fig. 4.1, we show the average point-to-point error between the fitted model and the manually-labelled shape, given different horizontal shape deviations. The flat area on the bottom of each curve indicates a stable (converged) output. At the right of Fig. 4.1, we give the convergence rate of two algorithms with different horizontal shape deviations.

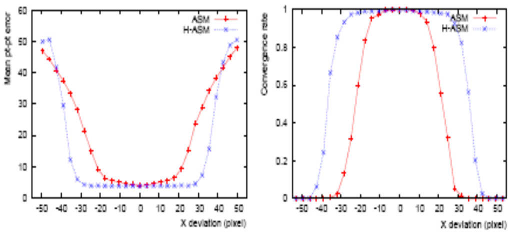

Figure 4.1: Convergence of the H-ASM algorithm as compared to ASM. Left: mean pt-pt error (in pixels) versus x-deviation; Right: convergence rate versus x-deviation.

4.1 Results of AFR using the adaptive Clustering Network Training

The training set contained 130 images of eight different faces with each face represented at a left facing (-90^0), straight ahead (0^0) and right facing (90^0) orientation. We presented the entire training set to the ACN over two repetitions to allow the cluster structure to stabilize. A learning rate of 0.25 was used to suppress any minor variations in the imagery.

5. Conclusions

In this papers, we discuss various topics in face recognition. These topics are divided into three key parts: face detection, facial feature extraction and face identification. For each part, we have reviewed the state-of-the art techniques in literature and proposed a set of novel techniques based on cascaded structures and designed a automatic face recognition system. During this system we make performance accuracy, efficiency and robustness.

We mainly focused on the algorithm design for face detection. The two-layer algorithm architecture for the face detector (ensemble and cascade) is not restricted to neural networks. It can be generalized to other embedded classifiers with classification *diversity* and complexity *scalability*.

The algorithm design for facial feature extraction. The first proposal we have developed is called H-ASM, which extends the Active Shape Model by using Haar-based local feature modelling.

Algorithms presented in this paper are efficient enough and of sufficiently high quality to initiate experimental embedded system implementations. Second, when looking to literature, it can be concluded that despite the many advances in face-recognition research, most of these advances are not suitable for consumer applications, because of the inherent complexity of the algorithms. For our work, we have taken *efficiency* and *accuracy* as a leading design principle from the beginning, and applied it to all stages of the face recognition. We have been able to make significant progress in improving the efficiency while maintaining the high performance. In this aspect, the work of this paper is certainly unique and can form an important basis for further embedded application design for face recognition.

References

[1] W.Y. Zhao, R. Chellappa, A. Rosenfeld, P.J. Phillips, Face Recognition: A Literature Survey, ACM Computing Surveys,Vol:35, 2003.

[2] G. Aggarwal, A. R. Chowdhury, R. Chellappa, A system identification approach for video-based face recognition, in: 17th International Conference on PatternRecognition, Vol. 4, 2004, pp. 175- 178.

[3] H.A. Rowley, S. Baluja, and T. Kanade, "Neural Network-Based Face Detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 1, pp. 23-38, Jan. 1998.

[4] R.L. Hsu, M. Abdel-Mottaleb, and A.K. Jain, "Face Detection in Color Images," Proc. Int'l Conf. Image Processing, pp. 1046-1049, 2001.

[5] M. Nakamura, H. Nomiya and K. Uehara, "Improvement of boosting algorithm by modifying the weighting rule", Annals of Mathematics and Artificial Intelligence, 41:95-109,2004.

[6] C.J. Edward, C.J. Taylor, and T.F. Cootes, "Learning to Identify and Track Faces in an Image Sequence," Proc. Int'l Conf. Automatic Face and Gesture Recognition, pp. 260-265, 1998.

[7] T. Cootes, G. Edwards and C.Taylor, "Active Appearance Models",IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, 2001, pp 681-685.

[8] Hyung-Soo Lee, Daijin Kim, Robust face tracking by integration of two separate trackers: Skin Color and facial shape, Pattern Recognition 40 (2007) ,pp:3225 - 3235

[9] A. Roy Chowdhury, R. Chellappa, R. Krishnamurthy and T.Vo, "3D Face Recostruction from Video Using A Generic Model", In Proc. of Int. Conf. on Multimedia and Expo, Lausanne, Switzerland, August 26-29, 2002.

[10] S. Baker and T. Kanade, "Limits on Super-Resolution and How to Break Them", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 9, September 2002, pp.1167-1183.2002.

[11] Abdenour Hadid; Matti Pietik?inen; Combining appearance and motion for face and gender recognition from videos, Pattern Recognition, Vol:42(11),2009,pp: 2818-2827.

[12] Freeman, W.T.; Tenenbaum, J.B., Learning bilinear models for two-factor problems in vision, 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1997, pp. 554 - 560.

[13] I. Matthews, R. Gross, and S. Baker, Appearance-based face recognition and light-fields, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 4, pp. 449-465,2004.

[14] A.J. O'Toole. *Handbook of Face Recognition*, chapter Psychological and neural perspectives on human face recognition. Springer, 2005.

[15] R. Koenen. Overview of the MPEG-4 standard, 2002. ISO/IEC JTC 1/SC 29/WG 11/N4668. [3] F. Zuo and P.H.N. de with. Automatic human face detection for a distributed video security system. In *Proceedings of the PROGRESS workshop*, pages 269–274, Oct. 2002.

S.Muralidharan **was** born in Kanchipuram,Tamil Nadu in 1981. He received the M.C.A. Degree from Thiruvalluvar university, Vellore, in 2006 First class with Distinction . He is currently doing Ph.D Degree in Dravidian University, Kuppam, A.P with the field of Image Processing.