Volume 2, No. 3, May-June 2011
ISSN No. 0976-5697
International Journal of Advanced Research in Computer Science
RESEARCH PAPER
Available Online at www.ijarcs.info

A Comprehensive Survey of Face Detection System Using Different Method

Megha Gupta*
Information and Technology department
Technocrats Institute of Technology
Bhopal, Madhya Pradesh., India
guptamegha_tct@yahoo.co.in

Prof. Neetesh Gupta
Information and Technology department
Technocrats Institute of Technology
Bhopal, Madhya Pradesh, India
gupta_neetesh81@yahoo.com

Abstract: In this paper we present a comprehensive and critical survey of face detection algorithms. Images containing faces are essential to intelligent vision-based human computer interaction, and research efforts in face processing include face recognition, face tracking, pose estimation, and expression recognition. However, many reported methods assume that the faces in an image or an image sequence have been identified and localized. To build fully automated systems that analyse the information contained in face images, robust and efficient face detection algorithms are required. Given a single image, the goal of face detection is to identify all image regions which contain a face regardless of its three-dimensional position, orientation, and lighting conditions. in this paper we have discuss different approach for face detection such as DWT fusion, image transformation, Age transformation and DCT based transform.

Keywords: Face Detection, DWT, DCT, Image Transformation

I. INTRODUCTION

The rapidly expanding research in face processing is based on the premise that information about a user's identity, state, and intent can be extracted from images, and that computers can then react accordingly, e.g., by observing a person's facial expression. In the last five years, face and facial expression recognition have attracted much attention though they have been studied for more than 20 years by psychophysicists, neuroscientists, and engineers. Many research demonstrations and commercial applications have been developed from these efforts. A first step of any face processing system is detecting the locations in images where faces are present. However, face detection from a single image is a challenging task because of variability in scale, location, orientation (up-right, rotated), and pose (frontal, profile). Facial expression, occlusion, and lighting conditions also change. Finding faces in an arbitrary scene and successfully recognizing them have been active topics in Computer Vision for decades. A general statement of the face recognition problem (in computer vision) can be formulated as follows: Given still or video images of a scene, identify or verify one or more persons in the scene using a stored database of faces. Although face reorganization and recognition is still an unsolved problem meaning there is no 100% accurate face reorganization and recognition system, however during the past decade, many methods and techniques have been gradually developed and applied to solve the problem. Basically, there are three types of methods in automatic face recognition: verification, identification and watch-list. In the verification method, a comparison of only two images is considered. The comparison is positive if the two images are matched. In the identification method, more than one comparison should be done to return the closest match of the input image. The watch-list method works similar to the identification method with a difference that the input face can also be rejected (no match). The method presented in this thesis consists of three steps: skin reorganization, face reorganization, and face recognition. The novelty of the proposed method is using a skin reorganization filter as a pre-processing step for face reorganization. A scheme of main tasks is shown in Figure 1

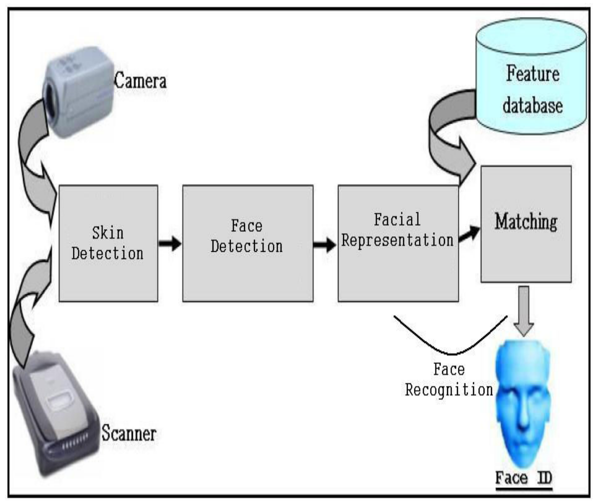

Figure (1). The rest of this paper is organized as follows. In section II Section III, Section IV, Section V and section VI conclusion.

II. IMAGE QUALITY ENHANCEMENT USING DWT FUSION

To improve the quality of both scanned and digital face images, we propose an image quality enhancement algorithm using DWT. In the proposed quality enhancement algorithm red channel (from RGB colour mode) and luma channel (i-e Y channel from YCbCr colour mode) are processed by applying multiscale retinex enhancement [1], wavelet denoising [2] and Wiener filtering [3]. In the watermarking literature, it is well established that red and luma channels are less sensitive to the visible (as well as perceptually invisible) watermarks. Since most of the document photographs have some form of watermarks, we, therefore select these two channels for enhancement. The output of these global enhancement algorithms are then fused using DWT. The proposed DWT algorithm identifies the good quality regions from each of the globally enhanced images and synergistically combines these enhanced quality regions to form a single composite image. This composite image contains feature-rich enhanced regions that are useful in face recognition. Details of the algorithm are as follows:

A. Let f be the colour face image to be enhanced. Let f^r and f^y be the red and luma channels respectively. These two channels are processed using multi-scale Retinex (MSR) algorithm [1], [4]. MSR is a modified form of the single scale retinex (SSR) that provides good dynamic range compression, colour in dependence from the spectral distribution of the scene illumination, and colour and lightness rendition. MSR is applied on both red and luma channels to obtain f^{rm} and f^{ym}.

B. Next step in the proposed quality enhancement algorithm is denoising of f^{rm} and f^{ym}. In this research, wavelet based adaptive soft thresholding scheme [2] is used for image denoising. The algorithm starts with computing generalized Gaussian distribution based soft threshold which is wavelet based denoising. After denoising f^{rm} and f^{ym}, we obtain f^{rm} and f^{ym} respectively.

C. After noise removal, genuine information that is useful in face recognition is blurred and therefore, it is important to apply a filter to deblur those edges. Experiments show that Wiener filter can restore the genuine facial edges. Applying Wiener filter on $f^{rm'}$ and $f^{ym'}$ produces f^1 and f^2.

D. After computing the globally enhanced red and luma channels. DWT fusion algorithm is applied on f^1 and f^2 to compute a feature rich enhanced face image, F. Single level DWT is applied on f^1 and f^2 to obtain the detail and approximation wavelet bands for these images. Let f^j_{LL}, f^j_{LH}, f^j_{HL}, f^j_{HH} be the four bands from face images and j= 1,2. To preserve the features of both the channels, coefficient from the approximation band of f^1 and f^2 averaged.

$$f^e_{LL} = \text{mean}(f^1_{LL}, f^2_{LL}) \quad (1)$$

Where f^e_{LL} is the approximation band of the enhanced image. All the three detailed sub bands are divided into windows of size 3X 3 ant the sum of absolute pixels in each window is calculated. For the I th window in the HL sub band of the two images, we select the window with the maximum absolute value and use it for f^e_{HL}. Similarly we obtain f^e_{LH} and f^e_{HH}. Finally, inverse DWT is applied on the four fused sub bands to generate a high quality face image.

$$F = I\,DWT(f^e_{LL}, f^e_{LH}, f^e_{HL}, f^e_{HH}) \quad (2)$$

This DWT fusion algorithm is applied on both scanned and digital face image.

III. DIGITAL IMAGE TO SCANNED IMAGE TRANSFORMATION

It was observed that, even after enhancement, the difference between the digital and scanned image of the same person is greater than the digital images of two different persons. To overcome this difference, we proposed a transformation algorithm that transforms the digital face image into a scanned like face image. An Eigen space is generated using the scanned face image and the digital face image is projected into this Eigen space for transformation [5]. The algorithm is describes as follows:

A. Let $F_D = [\vec{F}_{D1}, \vec{F}_{D2} \ldots \vec{F}_{Dn}]$ be the enhanced digital training images and $F_s = [\vec{F}_{S1}, \vec{F}_{S2} \ldots \vec{F}_{Sn}]$ be the enhanced scanned photo images, where n is the number of training images.

B. Average digital face image $\overrightarrow{aD}$ and average photo face image $\overrightarrow{aS}$ are computed for the digital and the photo training set respectively.

C. Eigen vector matrix E_D is computed from digital training set.

D. For transforming the probe image, average digital face $\overrightarrow{aD}$ is subtracted from the input digital face image $\vec{F}_{Dt}$

$$\vec{F}_{Dt'} = \vec{F}_{Dt} - \vec{a}_D \quad (3)$$

E. $\vec{F}_{Dt'}$ is projected in the Eigen space of $\vec{E}_D$ to compute the weight vector $\vec{w}_D$.

F. Digital face image is then reconstructed using the following equation,

$$\vec{F}^R_D = F_S V_D D_D^{-1/2} X \vec{w}_D \quad (4)$$

Where V_D represents the eigenvector matrix of digital face image training set and D_D represents the diagonal Eigen value matrix.

G. Average photo face $\vec{a}_S$ is added to the reconstructed digital face image $\vec{F}^T_D$,

$$\vec{F}^T_D = \vec{F}^R_D + \vec{a}_S \quad (5)$$

The proposed Eigen- transformation algorithm therefore transforms input digital face image $\vec{F}^{t'}_D$ to scanned face like image $\vec{F}^t_D$. The quality of reconstructed transformed digital face image is close to the scanned face image

IV. AGE TRANSFORMATION USING MUTUAL INFORMATION REGISTRATION

Once the face images are pre-processed and quality of the image is normalized, we minimize the age difference between gallery and probe face images. One way to address the challenge of facial aging is too regularly update the database with recent images or templates. However, this method is not feasible for applications such as border control and homeland security, missing persons and criminal investigations address this challenge, researchers have proposed several age simulation and modelling techniques. These technique model the facial growth that occurs over period of time to minimize the difference between probe and gallery images. Unlike, the conventional modelling approach, we proposed mutual information registration based age transformation algorithm to minimize the age difference between gallery and probe images. Mutual information is a concept from information theory in which statistical dependence is measured between two random variables.[6],[7]. Researchers in medical imaging have used mutual information based registration algorithms to effectively fuse images from different modalities such as CT and MRI [6], [7]. We have used this algorithm because there may be variations in the quality of pre-processed scanned and digital face images, and the registration algorithm should contend with these variations. Age difference minimization using registration of gallery and probe face images is described as follows:

Let F_G and F_P be the detected and quality enhanced gallery and probe face images to be matched. Mutual information between two face images can be represented as,

$$M(F_G, F_P) = H(F_G) + H(F_P) - H(F_G, F_P) \quad (6)$$

Where, H (.) is the entropy of the image and $H(F_G, F_P)$ is the joint entropy. Registering F_G with respect to F_P requires maximization of mutual information between F_G and F_P, thus maximizing the entropy $H(F_G)$ and $H(F_P)$, and minimizing the joint entropy $H(F_G, F_P)$. Mutual information based registration algorithms are sensitive to changes that occur in the distributions as a result of difference in overlap regions. To address this issue, Hill et al. [6] proposed normalized mutual information that can be represented as,

$$N\,M\,(F_G\;,F_P) = \frac{H(F_G)+H(F_P)}{H(F_{G,}F_P)} \qquad (7)$$

The registration is performed on the transformation space, T, such that

$$T = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ e & f & 1 \end{pmatrix}$$

Where a, b, c, d are shear, scale, and rotation parameters, and e, f are the translation parameters. Using the normalized mutual information and exploring the search space, T, we define a search strategy to find the transformation parameters, T *.

$$T^* = \arg\max \{N\,M\,(F_{P,}\,T\,(F_G))\} \qquad (9)$$

Gallery and probe face images (F_G and F_P) are thus registered using the transformation T *. This registered algorithm is linear in nature. To accommodate non linear variations in faces, multi resolution image pyramid scheme is applied which starts with building the Gaussian pyramid of both the gallery and probe images. Registration parameters are estimated at the coarsest level and used to warp the face images in the next level of the pyramid. The process is iteratively repeated through each level of the pyramid and a final transformed gallery face image is obtained at the finest pyramid level. In this manner, the global variations at the coarsest resolution level and local non-linear variations the finest resolution level are addressed. Once the age difference between the gallery and probe face image is minimized, face recognition algorithm can be efficiently applied to verify the identity of the probe image.

V. DISCRETE COSINE TRANSFORMED BASED LOCAL APPEARANCE MODELS

Local appearance face recognition is bases on statistical representations of the non- overlapping local facial regions and their combination at the feature level. The underlying idea is to utilize local information while preserving spatial relationships. In [8], the discrete cosine transform (DCT) is proposed to be used to represent the local regions. It has been shown to be a better representation method for modelling the local facial appearance compared to principal component analysis (PCA) and the discrete wavelet transform (DWT) in terms of face recognition performance. Feature extraction from depth images using local appearance-based face representation can be summarized as follows: The input depth image is divided into blocks of 8x8 pixels size. Each block is then represented by its DCT coefficients. These DCT coefficients are ordered using the zigzag scanning pattern [9]. From the ordered coefficients, M of them is selected according to the feature selection strategy, resulting in an M-dimensional local feature vector. Finally, the DCT coefficients extracted from each block are concatenated to construct the overall feature vector of the corresponding depth image. In order to compare the introduced local DCT-based representation with the depth representation, we calculated the ratio of within class variance with each representation on a training set which has also been used for identification experiments. We calculated the ratio of within class variance to between class variance for each representation unit and then averaged it over the representation units and the subjects. We obtained an average ratio of 0.5 with DCT-based local representation, and 0.67 with the depth representation. The lower ratio of within class variance to between class variance obtained by the proposed representation scheme indicates its better discrimination capability compared to the depth representation.

VI. CONCLUSIONS

This paper attempts to provide a comprehensive survey of research on face detection and to provide some structural categories for the methods described in over 10papers when appropriate, we have reported on the relative performance of methods. But, in so doing, we are cognizant that there is a lack of uniformity in how methods are evaluated and, so, it is imprudent to explicitly declare which methods indeed have the lowest error rates. In future we have used SVM classification for feature selection and improve the detection ratio of face.

VII. REFERENCES

[1]. Z. Rahman, D.J. Jobson, and G.A. Woodell, “Multiscale retinex for color image enhancement”, in Proceedings of IEEE International Conference on Image Processing, 1996, vol. 3, pp. 1003–1006.

[2]. S.G. Chang, Y. Bin, and M. Vetterli, “Adaptive wavelet thresholding for image denoising and compression”, IEEE TRANSACTIONS ON IMAGE PROCESSING, vol. 9, no. 1, pp. 1532–1546, 2000.

[3]. R.C. Gonzalez and R.E. Woods, Digital image processing, Prentice Hall, 2nd edition, 2002.

[4]. E.H. Land and J.J. McCann, “Lightness and retinex theory”, Journal of the Optical Society of America, vol. 61, no. 1, pp. 1–11, 1971.

[5] X. Tang and X. Wang, “Face sketch recognition”, IEEE Transactionson Circuits, Systems and Video Technology, vol. 14, no. 1, pp. 50–57,2004.

[6] D. Hill, C. Studholme, and D. Hawkes, “Voxel similarity measuresfor automated image registration”, in Proceedings of Third SPIEConference on Visualization in Biomedical Computing,1994, pp. 205–216.

[7] F. Maesa, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, “Multimodality image registration by maximization of mutual information”,IEEE Transactions on Medical Imaging, vol. 6, no. 2, pp.187–198, 1997.

[8] H. K. Ekenel and R. Stiefelhagen, “Local appearance-based face recognition using discrete cosine transform,” presented at the 13th Eur. Signal Processing Conf., Antalya, Turkey, 2005.

[9] R. C. Gonzales and R. E. Woods, Digital Image Processing. Upper Saddle River, NJ: Prentice-Hall, 2001.

