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

Computer Science and Engineering

Elixir Comp. Sci. & Engg. 55A (2013) 13251-13254

ISSN: 2229-712X

Face recognition under expressions and lighting variations using artificial intelligence and image synthesizing

C.JayaMohan, M.Saravana Deepak, M.L.Alphin Ezhil Manuel and D.C Joy Winnie Wise
Department of CSE, Alpha College of Engineering, Chennai, T.N, India.

ARTICLE INFO

Article history:
Received: 12 September 2012;
Received in revised form:
1 February 2013;
Accepted: 19 February 2013;

Keywords
Face recognition,
Constrained optical flow,
Artificial intelligence,
Synthesized image,
Masked synthesized image.

ABSTRACT
In this paper, we propose an integrated face recognition system that is robust against facial expressions by combining information from the computed intra-person optical flow and the synthesized face image in a probabilistic framework. Making recognition more reliable under uncontrolled lighting conditions is one of the most important challenges for practical face recognition systems. We tackle this by combining the strengths of robust illumination normalization. Our experimental results show that the proposed system improves the accuracy of face recognition from expressional face images and lighting variations.

Introduction

FACE recognition has been studied for several decades. Comprehensive reviews of the related works can be found in [14], [21]. Even though the 2-D face recognition methods have been actively studied in the past, there are still some inherent problems to be resolved for practical applications. It was shown that the recognition rate can drop dramatically when the head pose and illumination variations are too large, or when the face images involve expression variations. Pose, illumination, and expression variations are three essential issues to be dealt with in the research of face recognition. To date, there was not much research effort on overcoming the expression variation problem in face recognition, though a number of algorithms have been proposed to overcome the pose and illumination variation problems. To improve the face recognition accuracy, researchers have applied different dimension reduction techniques, including principle component analysis (PCA) [3], linear discriminant analysis (LDA) [13], independent component analysis (ICA) [1], discriminant common vector (DCV) [2], kernal- PCA, kernal-LDA [5], kernal-DCV [10], etc. In addition, several learning techniques have been used to train the classifiers for face recognition, such as SVM. Although applying an appropriate dimension reduction algorithm or a robust classification technique may yield more accurate recognition results, they usually require multiple training images for each subject. However, multiple training images per subject may not be available in practice.

This paper focuses mainly on the issue of robustness to expression and lighting variations. For example, a face Verification system for a portable device should be able to verify a client at any time (day or night) and in any place (indoors or outdoors). Traditional approaches for dealing with this issue can be broadly classified into three categories: appearance-based, normalization based, and feature-based methods. In direct appearance-based approaches, training examples are collected under different lighting conditions and directly (i.e. without undergoing any lighting preprocessing) used to learn a global model of the possible illumination variations.

The other category is to use optical flow to compute the face warping transformation. Optical flow has been used in the task of expression recognition [4], [8]. However, it is difficult to learn the local motion in the feature space to determine the expression change for each face, since different persons have expressions in different motion styles. Martinez [15] proposed a weighting method that independently weighs the local areas which are less sensitive to expressional changes. The intensity variations due to expression may mislead the calculation of optical flow. A precise motion estimation method was proposed in [14], which can be further applied for expression recognition. However, the proposed motion estimation did not consider intensity changes due to different expressions.

In this paper, we focus on the problem of face recognition from a single 2-D face image with facial expression. Note that this paper is not about facial expression recognition. For many practical face recognition problem settings, like using a passport photo for face identification at custom security or identifying a person from a photo on the ID card, it is infeasible to gather multiple training images for each subject, especially with different expressions. Therefore, our goal is to solve the expressive face recognition problem under the condition that the training database contains only neutral face images with one neutral face image per subject. In our previous work [11], we combined the advantages of the above two approaches: the unambiguous correspondence of feature point labeling and the flexible representation of optical flow computation. A constrained optical flow algorithm was proposed, which can deal with position movements and intensity changes at the same time when handling the corresponding feature.

Algorithm, we can calculate the expressional motions from each neutral faces in the database to the input test image, and estimate the likelihood of such a facial expression movement.

Tele:
E-mail addresses: alphacse138@yahoo.com

Using the optical flow information, neutral images in the database can be further warped to faces with the exact expression of input image. In this paper, we propose to exploit the two different types of information, i.e., the computed optical flow and the synthesized image, to improve the accuracy of face recognition. Experimental validation on the Binghamton University 3-D Face Expression (BU-3DFE) Database is given to show the improved performance of the proposed face recognition system. Since we do not attempt to solve the automatic facial landmark localization problem in this paper, the facial landmark points in our experiment are labeled manually.

Constrained Optical Flow Computation

The computational algorithms of traditional optical flow cannot guarantee that the computed optical flow corresponds to the exact pixels in different images, since the intensity variations due to expression may mislead the computation of optical flow. The brightness constancy constraint, however, is not valid in many circumstances. Therefore, with the generalized dynamic image model (GDIM) proposed by Negahdaripour and Yu [16], Teng et al. [19] generalized the optical flow constraint to

$$I_x(r)u(r) + I_y(r)v(r) + I_t(r) + m(r)I(r) + c(r) = 0 \qquad (1)$$

where $m(r)$ and $c(r)$ denote the multiplier and offset factors of the scene brightness variation field, I is the image intensity function, the subscripts x, y and t denote the spatiotemporal partial derivatives, r is a point in the spatiotemporal domain, and $[u(r),v(r)]^T$ is the motion vector at the point r .

Overall Face Recognition Processing Flow

Face recognition is a visual pattern recognition problem. A face is a three-dimensional object subject to varying illumination, pose, expression is to be identified based on its two-dimensional image (or three- dimensional images obtained by laser scan). A face recognition system generally consists of 4 modules - detection, alignment, feature extraction, and matching. Localization and normalization (face detection and alignment) are following processing steps before face recognition (facial feature extraction and matching) is performed.

The overall architecture of the face recognition processing flow system is shown in the Fig. 1.

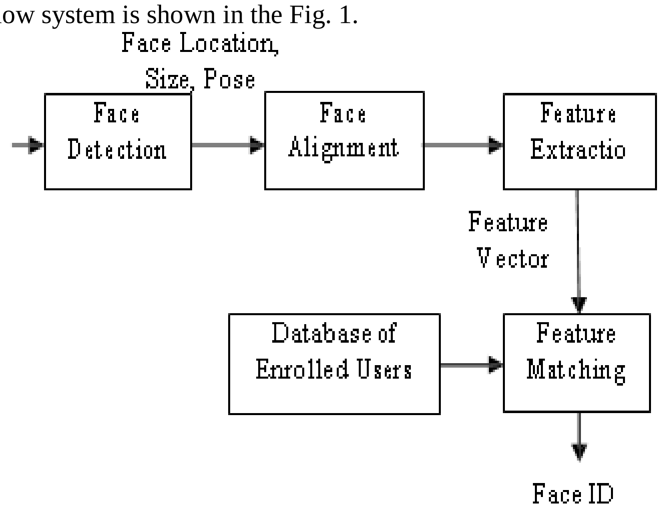

Fig. 1: Architecture of the Face Recognition Processing Flow System

The steps for above face recognition processing flow,Face detection segments the face areas from the background.

In the case of video, the detected faces may need to be tracked using a face tracking component.

Face alignment is aimed at achieving more accurate localization and at normalizing faces, whereas face detection provides coarse estimates of the location and scale of each face.

Facial components and facial outline are located; based on the location points,The input face image is normalized in respect to geometrical properties, such as size and pose, using geometrical transforms or morphing,The face is further normalized with respect to photometrical properties such as illumination and gray scale.

After a face is normalized, feature extraction is performed to provide effective information that is useful for distinguishing between faces of different persons and stable with respect to the geometrical and photometrical variations.

For face matching, the extracted feature vector of the input face is matched against those of enrolled faces in the database; it outputs the identity of the face when a match is found with sufficient confidence or indicates an unknown face otherwise.

Proposed Face Recognition System

Basic Concept

We treat the expression-invariant face recognition system as a probabilistic maximum a posteriori (MAP) classification problem. To do this, we formulate the problem as follows:

$$\arg\max_{N_i,E} P(N_i,E|I),\ i=1,2,\dots,N \qquad (2)$$

where I is the input image, N_i is the neutral face image for the ith subject in training data set, and E denotes the expression motion field between I and N_i. The optical flow field E is not specifically defined yet, which will be discussed later. The direction of E could be either from I to N_i or the opposite way. Based on the Bayes theorem and the assumption of independence Between N_i and E, (2) can be rewritten as

$$\arg\max_{N_i,E} P(N_i)P(E)P(I|N_i,E),\ i=1,2,\dots,N \qquad (3)$$

Face Recognition

Face recognition system that is robust against facial expressions by combining information from the computed intraperson optical flow and the synthesized face image. It is a computer application for automatically identifying (or) verifying a person from a digital image. One of the ways to do this is by comparing selected facial features from the image and a facial db. The database contains face models, each with a neutral face and six different expressions (angry, disgust, fear, happy, sad, and surprised) at different levels from level 1 (weakest) to 4 (strongest). Note that only the 2-D face images were used in our experiments. Fig. 2 shows the mask definition used for specifying the valid pixels in the face images. We first defined the standard mask image from the global neutral face image. A biometric system is generally a pattern recognition system which makes a personal identification by determining the authenticity of a specific physiological or behavioral characteristic indigenous from the user. Issues in the Face recognition is mainly due to following reasons: Identify similar faces (inter-class similarity), Accommodate intra- class variability due to head pose, illumination conditions, expressions ,facial accessories, aging effects, Cartoon faces.

Expression Images

The intraperson expression motion fields of the subjects in the training dataset, which are exclusive from the testing data, are collected in the training procedure. There are two advantages of the above optical flow normalization scheme: (1) all expressive face images of all subjects have the same dimension of motion fields, and (2) all optical flows are computed and represented with the same geometry.

Fig. 2 shows the mask definition used for specifying the valid pixels in the face images. We first defined the standard mask image [Fig. 2(b)] from the global neutral face image [Fig. 2(a)].

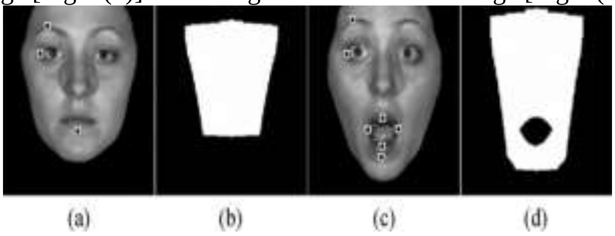

Fig. 2: Illustration of Mask Definition and Warping: (a) Reference Image and Feature Points, (b) Initial Mask, (c) Input Image, and (d) Warped and Sheared Mask for the Input Image

When there is an input image with expressions [Fig. 2(c)], the mask is then warped according to the three feature points shown in Fig. 2(a). Moreover, the region within the mouth is excluded in the region of interest [as shown in Fig.2(d)], since it cannot be synthesized due to the lack of texture in the corresponding region of the neutral image.

Lighting Variations

In lighting image, Normalization based approaches seek to reduce the image to a more "canonical" form in which the illumination variations are suppressed. Although a great improvement on raw gray values, their resistance to the complex illumination variations that occur in real-world face images is still quite limited. For example, even though LBP features are completely invariant to monotonic global gray-level transformations, their performance degrades significantly under changes of lighting direction and shadowing - see Fig. 3.

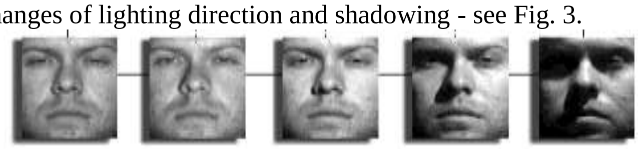

Fig. 3: Lighting Images

We propose an integrative framework that combines the strengths of all three of the above approaches. The overall process can be viewed as a pipeline consisting of image normalization, feature extraction and subspace representation.

Overall System Architecture

The overall architecture of the proposed face recognition system is shown in the Figure 4.

Experimental results in expression images and lighting variations

Our experiments were performed on a database containing 2-D face images, each with a neutral face and six different expressions (angry, disgust, fear, happy, sad, and surprised) at different levels, from level 1 (weakest) to 4 (strongest). Note that only the 2-D face images were used in our Experiments. Fig. 5 shows the 25 normalized face images of one subject after the normalization procedure.

Preprocessing

We manually labeled 21 feature points, including three points for each eyebrow and four points for each eye, one at the nose tip and the other six around the mouth region. With the labeled points, the distance between the outer corners of both eyes is used as the reference to normalize face images.

Mask Image

Using a mask image to eliminate the undesired area of an image is an important technique in remote sensing analysis. The principle behinds the mask technique is to multiply the source image by the mask image that contains two values—1 for preserved areas and 0 for undesired areas.

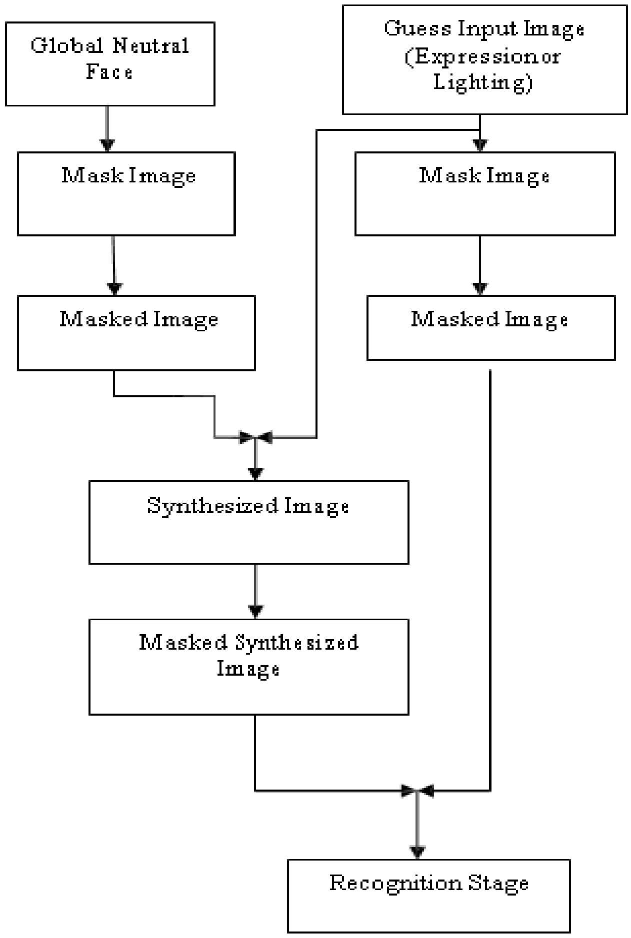

Fig. 4: Block Diagram of the Proposed Face Recognition System

Fig. 5: Sample Images for Expression. the Left-Top Most is the Neutral Face. the Others are the Face Images with Angry, Disgust, Fear, Happy, Sad, and Surprise Expressions in Columns From Left to Right with Increasing Levels in Rows From Top to Bottom

Fig. 6: Mask Image

Masked Image

Masking an image enables a developer to create images with irregular shapes dynamically. Masking is often used to create a user interface that is more compelling and less boring. The application of a mask to an input image produces an output image of the same size as the input.

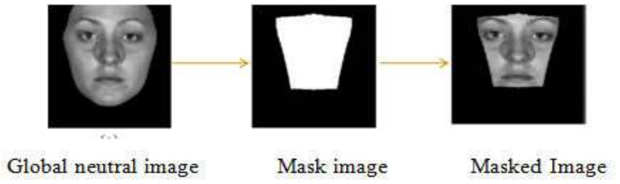

Fig. 7: Masked Images

Synthesized Image

Synthesis Image refers to synthesis of still images as well as synthesis of facial animations. For example, the technique of synthesizing facial expression images can be directly used for generating facial animations, and most of the facial animation systems involve the synthesis of still images. Digital Personnel is a computer-based facial expression synthesizer. It synthesizes animated, life-like, facial expressions of an individual in synchrony with that individual's speech. The system is speech driven, that is, as an individual speaks the appropriate facial expressions are generated simultaneously.

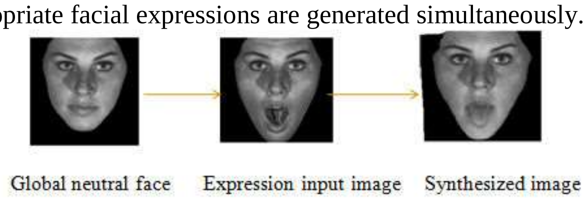

Fig. 8: Synthesized Images

Masked Synthesized Image

Masking is often used to create a user interface that is more compelling and less boring. The application of a mask to an input synthesized image produces an output image of the same size as the input.

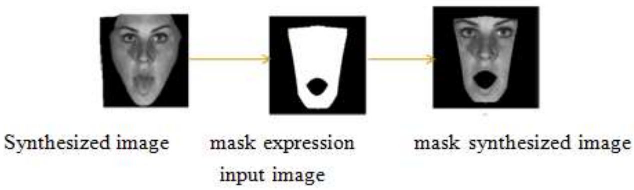

Fig. 9: Masked Synthesized Image

The same procedures have been used for face recognition under lighting variations.

Conclusion

In this paper, we proposed a 2-D expression-invariant face recognition system based on integrating the optical flow information and image synthesis. Only one neutral image for each candidate subject is needed in our face recognition system.

Two kinds of intraperson optical flow fields, u(x,y)@N0 and, u(x,y)@Ni were computed and used for expression motion likelihood calculation and expressive image synthesis, respectively. The proposed algorithm combines the face image MAP framework. As shown from the experimental results, the proposed face recognition system significantly improves the accuracy of face recognition on expressional face images and lighting variations.. However, the proposed integrated system is more computationally costly compared to the previous works, since the optical flow computation, image synthesis, and the probability calculations are needed for all candidates in the database. We have presented new methods for face recognition under uncontrolled lighting based on robust preprocessing and an extension of the Local Binary Pattern (LBP) local texture descriptor. There are following main contributions: (i) a simple, efficient image preprocessing chain whose practical recognition performance is comparable to or better than current (often much more complex) illumination normalization methods; (ii) a rich descriptor for local texture called Local Ternary Patterns (LTP) that generalizes LBP while fragmenting less under noise in uniform regions: (iii) a distance transform based similarity metric that captures the local structure and geometric variations of LBP/LTP face images better than the simple grids of histograms that are currently used.

References

[1] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, “Face recognition by independent component analysis,” IEEE Trans. Neural Netw., vol. 13, no. 6, pp. 1450–1464, Nov. 2002.

[2] H. Cevikalp, M. Neamtu, M. Wilkes, and A. Barkana, “Discriminative common vectors for face recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 1, pp. 4–13, Jan. 2005.

[3] M. A. Turk and A. P. Pentland, “Face recognition using Eigenfaces,”in Proc. IEEE Conf. Computer Vision and Pattern Recognition, Maui, HI, Jun.
1991, pp. 586–591.

[4] Y. Yacoob and L. S. Davis, “Recognizing human facial expressions from long image sequences using optical flow,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 18, no. 6, pp. 636–642, Jun. 1996.

[5] M.-H. Yang, “Kernel Eigenfaces vs. Kernel Fisherfaces: Face recognition using kernel methods,” in Proc. Int. Conf. Automatic Face and Gesture Recognition, Washington, DC, May 2002, pp. 215–220.

[6] T. Cootes, C. Taylor, D. Cooper, and J. Graham, “Active shape models—Their training and application,” Comput. Vis. Image Understand., vol. 61, pp. 18–23, 1995.

[7] T. Cootes, G. J. Edwards, and C. Taylor, “Active appearance models,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, pp. 681–685, 2001.

[8] I. A. Essa and A. Pentland, “A vision system for observing and extracting facial action parameters,” in Proc. IEEE Conf. Computer Vision Pattern Recognition, 1994, pp. 76–83.

[9] K. Fukunaga, Introduction to Statistical Pattern Recognition, 2nd Ed. New York: Academic, 1990.

[10]Y. H. He, L. Zhao, and C. R. Zou, “Kernel discriminative common vectors for face recognition,” in Proc. Int. Conf. Machine Learning and Cybernetics, Guangzhou, China, Aug. 18–21, 2005, pp. 4605–4610.

[11]C.-K. Hsieh, S.-H. Lai, and Y.-C. Chen, “Expression-invariant face recognition with accurate optical flow,” in Proc. PCM, Hong Kong, Dec. 11–14, 2007.

[12]X. Li, G. Mori, and H. Zhang, “Expression-invariant face recognition with expression classification,” presented at the 3rd Canadian Conf. Computer and Robot Vision, Jun. 2006.

[13]A. M. Martinez and A. C. Kak, “PCA versus LDA,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, no. 2, pp. 228–233, Feb. 2001.

[14]A. M. Martinez, “Matching expression variant faces,” Vis. Res., vol. 43, pp. 1047–1060, 2003.

[15]A. M. Martinez, “Recognizing expression variant faces from a single sample image per class,” presented at the IEEE Conf. Computer Vision and Pattern Recognition, Jun. 2003.