Computer Vision and Image Understanding 116 (2012) 1022–1031

Contents lists available at SciVerse ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Two dimensional synthetic face generation and verification using set estimation technique ☆

Madhura Datta [a], C.A. Murthy [b,*]

[a] UGC-Academic Staff College, University of Calcutta, 92 Acharya Prafulla Chandra Road, Kolkata 700 009, India
[b] Machine Intelligence Unit, Indian Statistical Institute, 203 B.T. Road, Kolkata 700 108, India

ARTICLE INFO

Article history:
Received 29 June 2010
Accepted 16 May 2012
Available online 6 June 2012

Keywords:
Feature extraction
Minimal spanning tree
Intra-class threshold
Set estimation
Inter class feature and intra class feature
PSNR

ABSTRACT

In this paper set estimation technique is applied for generation of 2D face images. The synthesis is done on the basis of inheriting features from inter and intra face classes in face space. Face images without artifacts and expressions are transformed to images with artifacts and expressions with the help of the developed methods. Most of the test images are generated using the proposed method. The measured PSNR values for the generated faces with respect to the training faces reflect the well accepted quality of the generated images. The generated faces are also classified properly to their respective face classes using nearest neighbor classifier. Validation of the method is demonstrated on AR and FIA datasets. Classification accuracy is increased when the new generated faces are added to the training set.

1. Introduction

Synthetic face generation for the purpose of face recognition has been explored in recent times. 2-Dimensional (2D)–3-dimensional (3D) reconstruction and generation of new faces with various shapes and appearances received attention of the biometric researchers. The popular solution to the problem is proposed by 2D and 3D modeling of faces. 3D models include face mesh frames, morphable models and depth map based models, where one needs to incorporate high quality graphics and complex animation algorithms. Flynn et al. [1] provided a survey of approaches and challenges in 3D and multi-modal 3D + 2D face recognition, Tao et al. [2] derived 3D head poses from 2D to 3D feature correspondences. Blanz et al. [3] proposed face recognition based on fitting a 3D morphable model with statistical texture. Four main approaches for 2D modeling are active appearance models (AAMs) [4], manifolds [5], geometry driven face synthesis methods [6] including face animation [7], and expression mapping techniques [8–10].

2D to 2D face reconstruction was first developed by Cootes et al. [4] and their active appearance model generated from the 2D face images is one of the powerful methods. Multi-view face reconstruction in 2D space is done by manifold analysis [5]. Geometry driven face synthesis by Zhang et al. [8], and expression mapping techniques are also useful in 2D face generation [11]. Table 1 provides comparative appraisal on advantages and disadvantages of some useful techniques of face synthesis. All the methods mentioned in Table 1, new face images are generated either from a model or from some functional properties.

In this paper the idea of developing new face images is based on statistical properties. A global formation of a face class for a particular person P can be assumed to be a set consisting of infinitely many face images. Face images of P differ from each other mainly because of muscle movements in different portions of the face. Different muscle movements result in different expressions. Movement of eyebrows, twitching of nose, muscle movement in cheeks, movement of lips, opening and closing of mouth along with different combinations these are some examples of movements in different parts of face.

If his/her face is continuously photographed, one may also note that, for any person, there exists a path joining those two images. Theoretically, a set $A \subseteq \mathcal{R}^m$ is said to be path connected, if for any two points $x, y \in A$, $\exists$ a continuous function f from [0, 1] to A such that $f(0) = x$, $f(1) = y$. That is, there exists a path containing infinitely many images joining x and y. Path connectivity is a valid assumption, since disconnected set of face images of the same person cannot occur. Therefore, we may mathematically define the face class as a path connected set. If we represent an image of a particular expression in the face of a person P by a vector x_0, then the set corresponding to the small variations in the same expression may be assumed to be a disc of radius $\varepsilon > 0$ around x_0. Using the set estimation principles, the face class for the person is estimated and the value of the radius $\varepsilon > 0$ from the estimated

☆ This paper has been recommended for acceptance by Martin Levine.
* Corresponding author.
E-mail address: murthy@isical.ac.in (C.A. Murthy).

1077-3142/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.cviu.2012.05.004

Table 1
Comparative advantages/disadvantages of 2D to 2D face synthesis with respect to some selected papers.

	Method	Advantages	Disadvantages
Cootes et al. [4]	Uses AAM and ASM models. Statistical texture and shape analysis techniques are used	Results are satisfactory	The initial feature points on face images are manually annotated in training stage
Huang et al. [5]	Faces in Manifold are reconstructed with expressions	Expressions are generated in neutral faces	Many face images of a person taken from different angles are needed for the construction of a manifold
Liu et al. [6] (ERI algorithm)	Expression ratio images are reconstructed	Applications are done on 2D faces and expressions are added successfully.	The method can only map expressions present in the probe image. Smooth transition from one face image to the other is not possible
Zhang et al.[8]	Facial expressions were synthesized feature point set	Expression editing software is available	Additional method is required for the generation of feature point set
Pyun et al. [12]	In this expression mapping technique, geometry controlled image warping method is used	Morphing can be done easily because of the knowledge of geometry of face	Geometrical properties of the face under consideration are to be found and stored
Pighin et al. [13]	Basis expression space is created for every person in the training set	Expressions of another person can be inherited	Construction of a person's expression space needs prior computations and also large memory space
Neely et al. [14]	Morphing operators are used for the construction of new faces	Easy to implement	Limited number of expressions are generated. Smooth transition from one face image to the other is not possible

set can be obtained. Inside this radius ε, say in m-dimensional Euclidean space, many different face images can be generated for the same person.

The method of set estimation is mainly used to find the pattern class and its multi-valued shape/boundary from its sample points in $\mathcal{R}^2$ [15–20]. Some investigation on estimation of α-hull for point sets in $\Re^3$ had been proposed by Edelsbrunner et al. [18]. Later Mandal et al. [19] extended the method to higher dimensions and found it useful in developing multi-valued recognition systems. As one can obtain the shape or boundary of a given finite set of points, the procedure of set estimation also generates the intuition for determination the radius of the disc. As the number of face images of the person increases, we shall be obtaining more information regarding the face class and hence the radius value needs to be decreased. Thus, the radius value is a function of the number of images and it is independent of the center of the disc. As a tool of set estimation, minimal spanning tree (MST) is used to calculate disc radius for each class. One of the methods of calculating the radius is (i) find MST of the set of n points with the edge weight as Euclidean distance, (ii) take L_n as the sum of the edge weights of MST and (iii) take the radius as, $\sqrt{l_n/n}$.

In contrast, in this paper, alternative procedure of using MST is adopted. Two algorithms are designed here to generate the new face images from the estimated sets. In the first method, new face images of a particular person are generated using features of that class only. In the second one, new face images are generated using features not only of that class, but also the features from other classes.

The face data bases used for experiment are AR [21] and FIA [22] databases. Both the AR and FIA data sets consist of frontal faces. In AR database, faces have illumination and artifact variations, while in the FIA database, the face images have varying facial expressions. Both the datasets have suitable properties for designing new morphed faces which are helpful for recognition purpose.

The paper consists of eight sections. Section 2 consists of brief statistical overview of set estimation technique and Section 3 introduces minimal spanning tree as set estimation tool. Two algorithms are designed for face generation in Section 4. Sections 5 and 6 consist of analysis of experimental results. In Section 7 classification accuracies of the method are discussed. Section 8 is the concluding section.

2. Brief statistical overview of set estimation for face classes

Before discussing the basic intuition for the applied method we assume that the dimensionality reduction technique has been already done using principal component analysis (PCA) [23,24]. Using PCA the image space is converted to face space. The whole process of set estimation, face class formation, and all the designed algorithms are carried out in reduced dimension, i.e. $\mathcal{R}^m$. Let a closed disc of radius r and centered at x_0 be represented by $\upsilon(x_0, r)$. That is, $\upsilon(x_0, r) = \{x : d(x_0, x) \leqslant r\}$, where d denotes the Euclidean distance.

If we represent a face image of an expression of a person P by a vector x_0, then the set corresponding to the small variations in the same expression may be assumed to be a disc of radius $\varepsilon > 0$ around x_0. The set corresponding to an expression of the person P may be taken as $\bigcup_{i=1}^{N}\{x \in \Re^m : d(x_i, x) \leqslant \varepsilon\}$, where $x_1, x_2, \ldots, x_N$ are N vectors corresponding to N images of the same expression for the same person. The dimension of the vectors is assumed to be m. The set corresponding to the union of all possible expressions of a person may also be taken as a connected set. The face class of a person is nothing but the set of all possible face images of that person.

In the above formulation, as the number of given face images of the person increases, we obtain more information regarding the face class and hence the radius value decreases. Thus, the radius value is a function of the number of images. Usually one may want to "estimate" a set on the basis of the given finitely many points. Grenander has formulated the set estimation problem as the problem of finding consistent estimate of a set [17].

Let α_n be an estimated set based upon the random vectors $X_1, X_2, \ldots, X_n$. Then α_n is said to be a consistent estimate of α, if $E_\alpha[\mu(\alpha_n \Delta \alpha)] \to 0$ as $n \to \infty$, where Δ denotes symmetric difference, μ is the Lebesgue measure and E_α denotes the "expectation" taken under α. Then

$$\alpha_n = \bigcup_{i=1}^{n}\{x \in \Re^2 : d(x, X_i) \leqslant \varepsilon_n\} \quad (1)$$

The discussed theorem did not mention a way of finding ε_n. Murthy [15] developed a way of finding ε_n for points in 2D spaces. He also generalized the method to any continuous density function on α, where α is a path connected set. His method has applications in different fields and a few of its modifications are all documented in literature [15–20]. None of these methods deals with generating new face images.

3. Minimal spanning tree as a tool of set estimation method [16]

Initially face space is created by any dimensionality reduction technique which is followed by the set estimation method. The