Multilinear Analysis based on Image Texture for Face Recognition

Hu-chuan Lu[1], Hao Chen[1]
[1]*Department of Electronic Engineering*
Dalian University of Technology
China 116023
E-mail: lhchuan@dlut.edu.cn

Yen-wei Chen[1,2]
[2]*College of Information Science and Engineering*
Ritsumeikan University
Japan 5258577
E-mail: chen@is.ritsumei.ac.jp

Abstract

In this paper, a multilinear approach based on image texture for face recognition is present. First, we extract the texture features of the facial images using the Local Binary Pattern (LBP) algorithm. Then, we apply the High-order Orthogonal Iteration (HOOI) algorithm, the algebra of higher-order tensors, to obtain a compact and effective representation of the facial images based on the texture features. Our representation yields improved facial recognition rates relative to standard eigenface and tensorface especially when the facial images are confronted by a variety of viewpoints and illuminations. To evaluate the validity of our approach, a series of experiments are performed on the CMU PIE facial databases.

1. Introduction

Facial images are formed by the interaction of multiple factors. People possess the remarkable ability to recognize faces confronted by a variety of expressions, poses and illuminations. Developing a robust computational model for face recognition remains a difficult problem.

Traditional face recognition adopts principal components analysis (PCA) [1, 2, 3], using the matrix singular value decomposition (SVD). The PCA approach calls for constrained conditions including frontal images and fixed illuminations. However, natural facial images result from the interaction of multiple factors related to facial expression, the pose of the head and the lighting conditions. These factors cause serious difficulties for conventional appearance-based face recognition approaches. High-order singular value decomposition (HOSVD) [4], the natural generalization of the matrix SVD, has been applied to face recognition. Vasilescu et al [5, 6] considers different factors such as viewpoints and illuminations in their facial image analysis and obtain better face recognition results than PCA.

In this paper, we employ a more powerful, tensor algebraic framework for the texture-based recognition of facial images that effectively deals with the multifactor variation. Our approach has the following characteristics: 1) the facial texture feature based on the LBP algorithm is an effective facial representation. Through dividing the original image into small regions, the computational cost in the learning stage is reduced owing to the reduced data dimensions.2) we can obtain the optimum truncated factor-specific mode matrix using the HOOI algorithm.3) the framework is robust to the variety of viewpoints and illuminations.

Following a review in Section 2 of the details of LBP algorithm and the multilinear algebra, Section 3 presents the framework of our multilinear analysis. In Section 4, we apply our framework to the recognition of the facial images. Section 5 concludes the paper.

2. Related Works

2.1. LBP Features

The original LBP operator was introduced by Ojala et al [7]. Later the operator [8] was extended to use neighborhood of different sizes using circular neighborhoods. The LBP operator $LBP_{P,R}$ produces 2^P different output values, corresponding to the different binary patterns that can be formed by the P pixels in the neighbor set with the radius R. It has been shown that certain bins contain more information than others

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

[8]. Ojala et al called these fundamental patterns uniform patterns denoted by the $LBP_{P,R}^{u2}$.

The LBP histogram contains information about the distribution of the local micro patterns. Face image can be seen as a composition of micro patterns which can be effectively described by the LBP features. However, the LBP histogram computed over the whole facial image considers the occurrences of the micro patterns without any indication about their locations. To encode shape information of faces, face images firstly are divided into small regions to extract LBP features separately (See Fig 1 for an illustration). The LBP features extracted from each sub-region are concatenated into a spatially enhanced texture feature histogram.

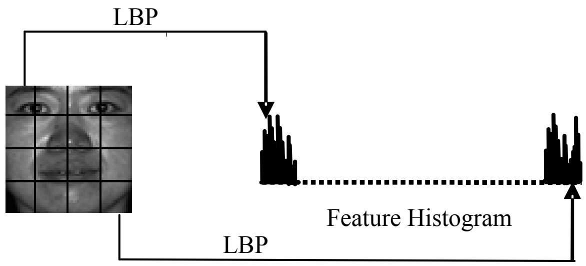

Figure 1. **Facial texture features based on LBP**

2.2. Relevant Tensor Algebra

A tensor, also known as n-way array or multidimensional matrix, is a higher order generalization of a vector (first order tensor) and a matrix (second order tensor). The mode-n vectors $A_{(n)}$ of an N order tensor A are the I_n-dimensional vectors obtained from A by varying index I_n while keeping the other indices fixed. The mode-n product $B=A\times_n U$ can be computed via the matrix multiplication $B_{(n)} = UA_{(n)}$, followed by a re-tensorization to undo the mode-n flattening.

Matrix SVD orthogonalizes the column and row space, the two associated spaces of a matrix. An order $N > 2$ tensor D is an N-dimensional matrix comprising N spaces. N-mode SVD is the generalization of conventional matrix SVD. It orthogonalizes these N spaces and decomposes the tensor as the mode-n product of N-orthogonal spaces. Thus a tensor can be expressed as a multilinear model as follows:

$$D=Z\times_1 U_1\times_2 U_2 \ldots\ldots \times_n U_n \ldots\ldots \times_N U_N \qquad (1)$$

Tensor *Z*, known as the core tensor, is analogous to the diagonal singular value matrix in conventional matrix SVD, but it does not have a simple, diagonal structure. The core tensor governs the interaction between the mode matrices $U_1 \ldots\ldots U_N$. Mode matrix U_n contains the orthonormal vectors spanning the column space of matrix $D_{(n)}$ resulting from the mode-n flattening of D.

The N-mode SVD algorithm for decomposing D (HOOI) according to equation (1) is as follows:

1. Initialize Un for n = 1……N using HOSVD
2. Repeat
   For n = 1……N do
   $Y=Z_1\times U^{(1)T} \ldots \times_{n-1} U^{(n-1)T} \times_{n+1} U^{(n+1)T} \ldots \times_N U^{(N)\,T}$
   Un: Rn leading left singular vectors of Y(n)
   end for
   Until fit ceases to improve
   $Z= D\times_1 U^{(1)\,T}\times_2 U^{(2)\,T}\times \ldots\ldots \times_N U^{(N)\,T}$
3. Return Z, U_1, U_2, ….., U_N

3. Overview of Our Framework

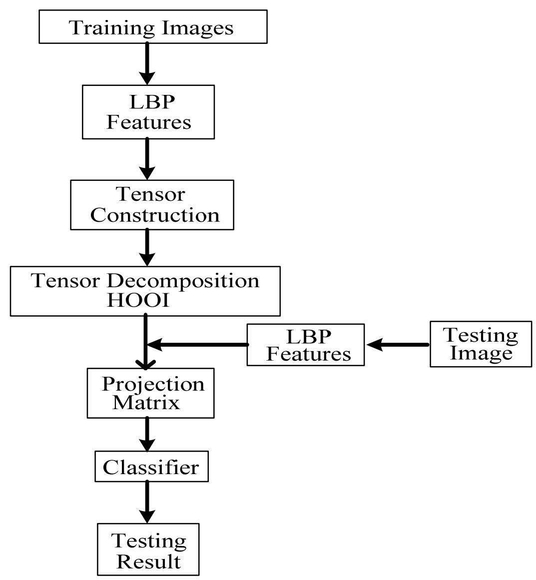

Figure 2. **Flow chart of our framework**

We illustrate our Framework (see Fig2 for an illustration) using a portion of the CMU PIE facial database: 40 subjects photographed in 5 viewpoints and 5 illuminations.

3.1. Facial Texture Features Extraction

Some parameters can be optimized for the LBP feature selection. The first one is the LBP operator, and another is the number of regions divided. For an illustration, we select the 59-bin $LBP_{8,1}^{u2}$ operator and divide the 64*64 pixels facial images into 16*16 pixels regions, giving a good trade-off between performance and feature vector length. Thus facial images are divided into 16(4*4) regions as shown in Fig 1, and represented by the LBP histogram features with length of 944(59*16).

3.2. Facial Texture Features Decomposition

Our tensor D based on image texture features is a four-order tensor. Applying multilinear analysis to D, using HOOI algorithm with N=4 we can obtain

$$D=Z\times_1 Upeople\times_2 Uviews\times_3 Uillums\times_4 Ufeatures \quad (2)$$

Where the core tensor Z governs the interaction between the factors represented in the 4 mode matrices: The mode matrix Upeople spans the space of people parameters, the mode matrix Uviews spans the space of viewpoint parameters, the mode matrix Uillums spans the space of illumination parameters and the mode matrix Ufeatures orthonormally spans the space of image texture features.

The facial image database comprises 25 images per person that vary with viewpoints and illuminations. PCA represents each person as a set of 25 vector-valued coefficients. The length of each coefficient vector is 40*5*5=1000. By contrast, multilinear analysis enables us to represent each person, regardless of viewpoints and illuminations, with the same coefficient vector of dimension 40 relative to the bases comprising the 40*5*5*944 tensor.

$$B = Z \times_2 Uviews \times_3 Uillums \times_4 Ufeatures \quad (3)$$

This many-to-one mapping is useful for face recognition. Each image is represented with a set of coefficient vectors representing the person, viewpoint and illumination factors.

3.3. Face Recognition

We can extract the mode matrix Upeople which contains row vectors c_p^T of coefficients for each person p and constructs the basis tensor B through the tensor decomposition. We index into the basis tensor for a particular viewpoint v and illumination i to obtain a subtensor Bv,i of dimensions 40*1*1*944. We flatten Bv,i along the people mode to obtain the 40*944 matrix. $B_{v,i\,(people)}$. Note that a specific training image d of person p in viewpoint v and illumination i can be written as

$$c_{v,i} = B_{v,i(people)}^{-T} d \quad (4)$$

Now, given an unknown facial image I, we use the projection operator to project I into a set of candidate coefficient vectors $c_{v,i} = B_{v,i(people)}^{-T} I$ for every v and i. Our recognition algorithm compares each $c_{v,i}$ against the person-specific coefficient vectors C_p. The best matching vector C_p which yields the smallest value of $\| c_{v,i} - c_p \|$ among all viewpoints and illumination identifies the unknown image I as the person p.

4. Experiments and Analysis

The CMU Pose, Illumination, and Expression (PIE) database contains more than 40000 facial images of 68 people. The images are acquired over different poses, under variable illumination conditions. In the experiments below, all the images are grayscale to a resolution of 64*64 pixels and two sub databases are used to evaluate our proposed approach.

Database1: we train our model on an ensemble comprising image of 40 people, captured from 3 viewpoints indexed as C11, C27 and C37, with 5 illuminations indexed as 07,08,09,10 and 13. We test our model on other images in the same 40 person dataset acquired from 2 different viewpoints indexed as C05 and C29 under the same 5 illuminations.

Database2: We train our model on an ensemble comprising image of 40 people, captured from 3 viewpoints index as C11, C27 and C37, with 3 illuminations indexed as 08, 10 and 13. We test our model on other images in the same dataset acquired from 2 different viewpoints indexed C05 and C29 under 2 different illuminations indexed as 07 and 09.

4.1 The Effect of Different Window Sizes

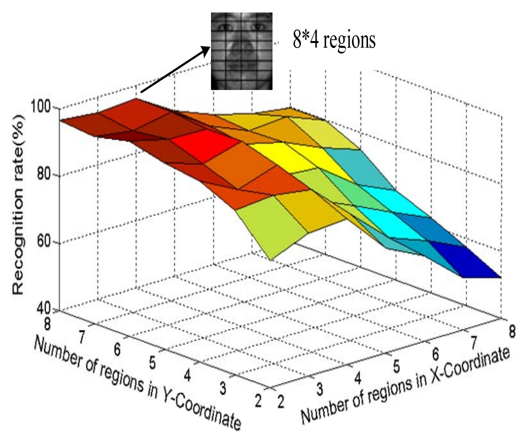

Figure 3. **Recognition rate (%) in different window sizes**

We use Database1 to show the effect of different window sizes. The recognition rate for $LBP_{8,1}^{u2}$ as a function of the window size is plotted in Figure 3. The original 64*64 pixel image is divided into i*j windows, i=2…8 and j=2…8 resulting in window sizes from 32*32 to 8*8. The small windows are not tested because of the high dimensions of the feature vector. The 8*16 pixel window is selected since it is a good trade-off between recognition performance and feature vector length.

4.2. The Effect of Dimensionality Reduction in Factor-specific Mode Matrix

We use Database1 to show the effect of truncated mode matrix. We apply HOOI algorithm to the dimensionality reduction of factor–specific mode matrix. The recognition rate for different truncated mode matrix Uviews and Uillumis is plotted in Fig 4 below. We select the truncated mode matrix Uviews(3*2) and Uillumis(5*2) because of the better recognition performance.

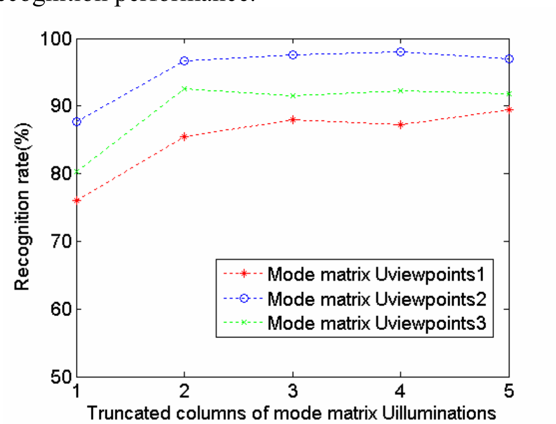

Figure 4. **Recognition rate (%) using the truncated factor-specific mode matrix**

4.3. Recognition Accuracy Comparison

Table 1. **Recognition accuracy (%) comparison under variable viewpoints on Database1**

Algorithm	Accuracy
EigenFace	53.25
TensorFace	56.00
LBP+PCA	85.75
LBP+Tensor	96.25

Table 2. **Recognition accuracy (%) comparison under variable viewpoints and illuminations on Database2**

Algorithm	Accuracy
EigenFace	21.25
TensorFace	44.37
LBP+PCA	75.62
LBP+Tensor	93.75

For Database1, we extract the texture features using the $LBP_{8,1}^{u2}$ operator with the 8*16 pixel window and apply HOOI algorithm to obtain the truncated mode matrix Uviews(3*2) and Uillumis(5*2). For Database2, we extract the texture features using the $LBP_{8,1}^{u2}$ operator with the 8*16 pixel window and apply HOOI algorithm to obtain the truncated mode matrix Uviews (3*2) and Uillumis (3*2).From the experimental result, we can observe that our algorithm (LBP+Tensor) outperforms the other algorithms especially when the facial images are confronted by a variety of viewpoints and illuminations. The experiments show that our framework is robust to viewpoints and illuminations.

5. Conclusions

In this paper, we have presented a multilinear approach based on image texture for face recognition. First, the texture features of facial images are extracted by the LBP algorithm. The extracted texture features can not only represent the facial images effectively, but also reduce the dimensions of original data. Then the HOOI algorithm is applied to facial analysis and recognition. In addition, we can obtain the optimized parameters for the feature selection and the optimum and truncated factor-specific mode matrix through the experiments. The experimental results on CMU PIE databases show that our approach outperforms the EigenFace and TensorFace.

References

[1] L. Sirovich and M. Kirby. Low dimensional procedure for the characterization of human faces [J]. *Journal of the Optical Society of America A.*, 4:519–524, 1987.

[2] M. Kirby and L. Sirovich. Application of the KL Procedure for the Characterization of Human Faces [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.12, no.1, pp.103-108, Jan.1990.

[3] M. A. Turk and A. P. Pentland. Eigenfaces for recognition [J]. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.

[4] L. Lathauwer, B. Moor and J. Vandewalle. A multilinear singular value decomposition [J].*Matrix Analysis and Applications*, 21(4):1253－1278, 2000.

[5] M. A. O. Vasilescu and D. Terzopoulos. Multilinear Analysis of Image Ensembles: TensorFaces [C]. ECCV 2002, pp. 447-460.

[6] M.A.O. Vasilescu and D. Terzopoulos, Multilinear Image Analysis for Face Recognition *Proc. 16th Int'l Conf. Pattern Recognition,* pp. 511-514, 2002.

[7] T. Ojala, M Pietikinen, and D. Harwood. A comparative study of texture measures with classification based on featured distribution [J]. *Pattern Recognition*, 1996, 29(1):51-59.

[8] T. Ojala, M. Pietikinen, and T. Menp. Multiresolution grayscale and rotation invariant texture classification with local binary patterns [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(7):971-987.