Two-dimensional Heteroscedastic Linear Discriminant Analysis for Age-group Classification

Kazuya Ueki†‡, Teruhide Hayashida†, and Tetsunori Kobayashi†
† Science and Engineering, Waseda University Okubo 3–4–1, Shinjuku-ku, Tokyo, 169–8555 Japan
‡ NEC Soft, Ltd. Shinkiba 1–18–7, Koto-ku, Tokyo, 136–8627 Japan

Abstract

This paper presents a novel LDA algorithm named 2DHLDA (2-Dimensional Heteroscedastic Linear Discriminant Analysis). The proposed algorithms are applied on age-group classification using facial images under various lighting conditions. 2DHLDA significantly overcomes the singularity problem, so-called ‘Small Sample Size’ problem (S3 problem), and the original feature space is split into useful dimensions and nuisance dimensions to reduce the influence of different lighting conditions. A two-phased dimensional reduction step, namely 2DHLDA+LDA, is used in our experiment. Our experimental results show that the new 2DHLDA-based approach improves classification accuracy more than the conventional 1D and 2D-based approaches.

1. Introduction

We have been researching and developing an automatic gender and age-group classification system by extracting human features from images. For over the last 20 years, gender classification from facial images has been one of the most actively researched topics in this field. Two types of approaches have been proposed to detect facial features: (i) geometry-based and (ii) appearance-based approaches. Geometry-based methods rely on metric features such as face width, eye size, mouth size, nose size, eyebrow thickness and distances between salient feature points (eyes, nose, mouth, etc.)[1][7]. Appearance-based methods find the decision boundary between male and female classes from training images without extracting any geometric features, and train classifiers such as neural network[2][9], principal component analysis (PCA)[3] and support vector machine (SVM)[14].

On the contrary, very few attempts have been made at age-group classification[10][13]. Even with the human eye, estimates of a candidate’s age are often inaccurate. One of the reasons why age-group classification is difficult is that enormous time and expense is required for collecting images including a wide variety of age-groups under the same lighting conditions, due to privacy and portrait rights. Our database, WIT-DB (Waseda human-computer Interaction Technology - DataBase), contains images from a wide variety of lighting conditions and age-groups. In addition, the number of images under a particular lighting condition is unbalanced. To achieve accuracy in classification, a new framework that can reduce the influence of different lighting conditions, is essential before classification. In this paper, our goal is to find the projection method that can reduce the various lighting conditions for feature extraction.

PCA and linear discriminant analysis (LDA) are commonly used for feature extraction in face recognition. LDA can enhance class separability of all sample images for classification purposes, but whenever the number of samples is less than the dimensionality of the samples, the execution of LDA may encounter the so-called Small Sample Size Problem (S3 Problem), therefore the transformation matrix can not be computed. The S3 Problem is often encountered when we use facial images in face recognition because of the high dimensionality.

Due to the S3 problem, before LDA can be applied to reduce dimensionality, PCA is commonly used for dimensionality reduction: PCA+LDA[4][6]. However, PCA step may extract nuisance dimensions such as lighting conditions, and degenerate classification accuracy because of discarding important discriminative dimensions. Therefore, direct LDA (DLDA) methods, which can accept a small number of high-dimensional data and optimizes Fisher’s criterion directly without dimensionality reduction steps, were proposed to solve the S3 problem[11][12]. Recently, 2D-based approaches; 2-dimensional principal component analysis (2DPCA)[15] and 2-dimensional linear discriminant analysis (2DLDA)[16][17], which extract features directly from 2-D images without a vectorization procedure, were proposed to reduce the computational cost and avoid the singularity problem.

Incidentally, Heteroscedastic Linear Discriminant Anal-
ysis (HLDA) is becoming popular in state-of-the-art speech recognition systems. HLDA can separate the original feature space into two independent subspaces; useful dimensions and nuisance dimensions. However, calculating the transformation matrix is difficult due to high dimensionality and extreme sparseness of the data.

In this paper, a new method 2-dimensional heteroscedastic linear discriminant analysis (2DHLDA), which is based on 2D-based methods and HLDA algorithm, is discussed in order to find the most discriminant projection vectors for age-group classification. This method reduces overcomes the singularity problem implicitly, and reduces the influence of different lighting variations.

We introduce conventional 2D-based projection methods in Section 2, and our proposed framework in Section 3. Experimental results for our method and conventional methods are discussed in Section 4, and conclusions are presented in Section 5.

2. Review of Previous 2D-based Approaches

As facial images have very high dimensionality, we need efficient feature extraction before classification in order to robustly classify age-groups under various lighting conditions with speed. In this section, conventional 2D-based projection methods are introduced, namely 2DPCA[15] and 2DLDA[16][17]. They have been developed for the projection from high dimensional facial space to the lower dimensional space.

2.1. 2DPCA

2DPCA[15] projects an $h \times w$ random image matrix $X$ onto $w \times \tilde{w}$ transformation matrix $W_{2DPCA}$,

$$ Y = XW_{2DPCA}, $$

where $Y$ denotes a $h \times \tilde{w}$ feature matrix. Transformation matrix $W_{2DPCA}$ is calculated by solving the eigenvalue problem of image covariance (scatter) matrix $G$:

$$ G = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^T (X_i - \bar{X}), $$

where $n$ denotes the total number of training data.

2.2. 2DLDA

2DLDA[16][17] uses the within-class scatter matrix $S_w$ and the between-class scatter matrix $S_b$ as well as in LDA:

$$ S_w = \frac{1}{n} \sum_{j=1}^{c} \sum_{X \in c_j} (X - \bar{X}_j)^T (X - \bar{X}_j), $$

where $c$ denotes the number of classes, $n_j$ denotes the number of samples in class $c_j (j = 1, 2, \cdots, c)$, $X$ denotes an $h \times w$ random image matrix, $\bar{X}$ denotes the average of all training samples $X_i$, and $\bar{X}_j$ denotes the average of samples in $c_j$-th class. One way to find the transformation matrix $W_{LDA}$ is to use Fisher’s criterion. $W_{LDA}$ can be constructed by the set of largest eigenvalues of $S_b S_w^{-1}$. Projecting $X$ onto $W_{2DLDA}$ yields an $h \times \tilde{w}$ feature matrix $Y = XW_{2DLDA}$.

3. Proposed Method

3.1. 2DHLDA

HLDA[8], viewed as a generalization of LDA, tries to find the best linear discriminant, but it differs from LDA in the underlying assumptions. HLDA removes the restriction that all the within-class covariance matrices are the same. We propose 2DHLDA, which is an extension of the 2DLDA method and assumes the covariance matrices to be different for all classes as well as in the case of HLDA. The 2DHLDA projection matrix $W$ is written as

$$ Y = XW = X \left[ W_{\tilde{w}} W_{w-\tilde{w}} \right], $$

where $W_{\tilde{w}}$ is a matrix consisting of the first $\tilde{w}$ of matrix $W$ and $W_{w-\tilde{w}}$ consists of the remaining $w - \tilde{w}$ rows, the top $\tilde{w}$ dimensions contain discriminatory information, the useful dimensions, and the final $w - \tilde{w}$ dimensions contain the nuisance dimensions.

The optimal transformation matrix is calculated by maximizing the log-likelihood Gaussian function in [8]. The final solution can be obtained as

$$ \hat{W} = \arg \max \left\{ -\frac{n}{2} \log |W_{w-\tilde{w}}^T S_t W_{w-\tilde{w}}| - \sum_{j=1}^{c} \frac{n_j}{2} \log |W_{w-\tilde{w}} S_{w}^{(j)} W_{w-\tilde{w}}| + n \log |W| \right\}, $$

where $\hat{W}$ is the estimate of the parameter $W$, $S_t = S_w + S_b$, and $S_{w}^{(j)}$ denotes the scatter matrix in $c_j$-th class.

3.2. 2DHLDA+LDA

In this paper, a two-phased approach 2DHLDA+LDA is employed: 2DHLDA is done for the first dimensionality reduction step from $32 \times 32$ to $32 \times 10$, and LDA is used for the second dimensionality reduction from 320 dimensions to 4 dimensions.
One advantage of the 2DHLDA+LDA approach is that it can solve the S3 problem, since the transformation matrix is computable using a smaller amount of data as compared to LDA. For example, in the case of 32×32 monochrome images, when we derive 10 component vectors (32×10 features in total) using 2DHLDA and the dimensionality is reduced from 320 to 4 using LDA, 320 images are sufficient, whereas more than 1,024 images are necessary using LDA only approach.

### 4. Experiments for Age-group Classification

#### 4.1. WIT-DB

We developed WIT-DB using images taken from approximately 5,500 different Japanese subjects (about 2,500 females and about 3,000 males) with 1 to 14 image samples from each subject. WIT-DB contains un-occluded frontal facial images, and most of the facial expressions in these images are normal except for smiles in some of the images. Table 1 shows the number of data in each age-group class. The image size used in this paper is a facial region of 32×32 pixels, as shown in Figure 1. These images were taken from subjects in a wide variety of lighting conditions and age-groups, from 3 to 85 years of age, because our research is motivated by real-world application that must be robust against lighting variations. Our database has 11-class age-groups registered. These age-groups are based on actual age, and not appearance age, as shown in Table 1. Our goal is to classify 11-class age-groups with high accuracy.

#### 4.2. Outline of Experiments

Histogram equalization is done to each image to compensate for the lighting conditions and improve the contrast of images. data from 12,008 female and 14,214 male images has been individually treated in this experiment, and the performance of age-group classification is evaluated using 2-fold cross validation in each gender.

We use the new proposed method, 2DHLDA+LDA, in order to reduce dimensionality for feature extraction, and also use LDA, 2DPCA+LDA and 2DLDA+LDA for comparison purposes.

After the process of two-phased projection, eleven Gaussian models from each age-group which are in Table 1, are constructed for each gender by using the EM algorithm. Then, the likelihood scores for each class are computed and the class with the highest likelihood is chosen. For the reason of the difficulty in age-group classification, the classification rate in the 10-year range, which includes the contiguous class with the higher likelihood, and in the 15-year range, which includes both contiguous classes, are observed.

#### 4.3. Experimental Results

Table 2 shows classification accuracy using the LDA only, 2DPCA+LDA, 2DLDA+LDA and 2DHLDA+LDA approaches, when age-groups are in the 5-year, 10-year, and 15-year range respectively. For the LDA only approach, the first 4 dimensions are selected. For the 2DPCA+LDA, 2DLDA+LDA and 2DHLDA+LDA approaches, dimensionality is reduced from 32×32 to 32×10 in the first stage, and from 320 to 4 in the second stage. The average of all training samples $\bar{X}$ is set to 0 in our experiment. These parameters are set experimentally.

Experimental results verify high efficiency of our approach, 2DHLDA+LDA. In every method, classification rates in male are higher than the ones in female, and the result shows classifying age-group using female images to be difficult.

Moreover, Figure 2 shows classification accuracy of age-group using 2DHLDA+LDA. The row denotes the class based on the actual age, and the column denotes the class based on experiment, and the class with the higher classification rate is darker in color. In terms of younger age-groups (under 19) and older age-groups (over 55), classification rates are higher in each gender, however as for age-groups between 20 and 49, especially in females, classification decreases as shown in the figure.
Table 2. The classification rate of age-group based on different projection methods.

<table>
<thead>
<tr>
<th>approaches</th>
<th>female</th>
<th>male</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5-year</td>
<td>10-year</td>
</tr>
<tr>
<td>LDA</td>
<td>40.00%</td>
<td>59.46%</td>
</tr>
<tr>
<td>2DPCA+LDA</td>
<td>42.76%</td>
<td>62.28%</td>
</tr>
<tr>
<td>2DLDA+LDA</td>
<td>43.05%</td>
<td>62.21%</td>
</tr>
<tr>
<td>2HDLDA+LDA</td>
<td>43.30%</td>
<td>63.51%</td>
</tr>
</tbody>
</table>

Figure 2. The confusion matrix of age-group classification using female data (left) and male data (right).

5. Conclusion

In this paper, we proposed a two-phased approach (2DHLDA+LDA) for age-group classification using facial images under various lighting conditions. Our approach does not require PCA, which extracts lighting condition variations, and solves the S3 problem under a small amount of samples. Additionally our experiments show that the 2DHLDA+LDA approach is superior to conventional 1D-based and 2D-based approaches. Thus, effective feature extraction, not for lighting condition variation but for the age-group classification, is achieved.

For future work, we plan to solve some familiar problems such as viewing orientation, partial occlusion of facial features and facial expression. Moreover, we plan to evaluate the classification rates using images under unknown lighting conditions, and find the difference between actual age and appearance age.

References