Abstract

Feature design and feature selection are two key problems in facial image based age perception. In this paper, we proposed to using ranking model to do feature selection on the haar-like features. In order to build the pairwise samples for the ranking model, age sequences are organized by personal aging pattern within each subject. The pairwise samples are extracted from the sequence of each subject. Therefore, the order information is intuitively contained in the pairwise data. Ranking model is used to select the discriminative features based on the pairwise data. The combination of the ranking model and personal aging pattern are powerful to select the discriminative features for age estimation. Based on the selected features, different kinds of regression models are used to build prediction models. The experiment results show the performance of our method is comparable to the state-of-art works.

1 Introduction

Due to the various applications of face image analysis, face identity, facial pose estimation, facial expression analysis and gender classification have been widely studied in the literature of computer vision. However, facial image based age estimation has not been extensively explored yet due to the difficulty of data collection. This problem is now partly alleviated due to the public dissemination of the FG-NET Aging Database[1]. In recent years, automatic age estimation has attracted more research interests, because of its wide potential applications, such as specific human computer interfaces, and improving recognition system efficiency. Although shape feature as ASM and texture feature have got some success on age estimation problem[8][13], developing and selecting efficient and discriminative features is still a key problem. In this paper, we proposed to use ranking model to select powerful discriminative features.

1.1 Previous Work

Age estimation is one typical regression problem, however, due to the difficulty of data collection, preliminary works [3][11][10][19] formulate age estimation as a classification problem. These works grouped adults and children, and split adults into young adults and seniors. The method AGES proposed by Xin[8] defined an image sequence of one subject as an aging pattern, which is represented using PCA model, and age estimation is performed by searching the proper position at age patterns. Although AGES got great success in age estimation, it heavily depends on the age distribution of training set. If there is no training sample at one specific age, intuitively, it’s impossible to correctly predict the testing instance at the corresponding age. Recently, researchers attempt to conduct more precise age estimation, and they try to predict the facial age directly by using different kinds of regression models [16][7]. Yun [7] learned low dimensional aging manifold from image intensities and age estimation is formulated as a regression problem in manifold space. Yan [16] proposed auto-structured regressor based on nonnegative labels. In a summary, these above methods focus on regression model design and ignore how to select powerful features.

1.2 Motivation

Geometric features and texture features are widely used in facial image analysis, active appearance model(AAM)[4] is the most popular face model used to project face images onto a low dimension parametric space. Xin[8] and Lanitis [12] built their model to estimate age based on AAM features. Recently, Suo[13] designed sparse features which also include AAM features, and hierarchical face model to predict age. The encouraged experiment results in [13] show the powerful features could give a lot of help on age estimation. As we know, AAM heavily depends on the face alignment result, however, facial image quality is a big problem for alignment in the practical application. Additionally, AAM is usually not expressive enough to represent...
details critical to age perception, and as an important cue for age perception.

Although Suo [13] has presented powerful features on age prediction, their method requires precise alignment which is also not reliable in the practical system. Eye localization is much more robust than other facial mark points. Therefore, the texture feature extracted from normalized face by eye’s position is more reliable. Texture features such as local binary pattern was used by Yang [19] to do age group classification and achieves good experiment result. Due to the large amount of raw feature size, Yang[19] used adaboost to combine weak classifiers, this procedure equals to feature selection when one weak classifier is built on the corresponding feature. With the prior knowledge, Suo [13] directly took informative parameters for age estimation as explicit features and some discriminative texture features were used as implicit features. In a word, [19] and [13] selected features by the facial difference between different age groups, the aging face variance within individual subject is ignored. In this paper, we try to explore the discriminative features within the subject to describe the age variance. Recently, ranking methods are used in face recognition [2], face alignment[15], and facial expression analysis [18]. Inspired by the above ideas in[2][15], we extend ranking method on age prediction and use ranking model to do feature selection. The paper is organized as following: the haar features are shortly introduced in section 2. Section 3 presents the ranking model, experiment is shown in section 4 followed by the conclusion in section 5.

2 Features

We believe simple texture features could also get promising performance with proper strategy [19][17]. Therefore, in this paper, we use the haar-like features to represent facial appearance due to its good properties and low computation cost, especially in facial detection [14], face alignment[15] and facial expression recognition[17]. In our experiment, the haar-like feature descriptors are directly from the OpenCV. The minimum rectangle is 2 by 2, and the largest one is $\frac{1}{2}n \times \frac{1}{2}n$, where $n$ is the image width. Different scale rectangles slides through the image to get thousands of haar-like features. We resize the facial images with the size of 64 $\times$ 64, so totally we get 195552 haar-like features in one face image. Due to thousands of haar features are exhaustively extracted from one face image, this leads to high dimensional feature vectors and causing the training process time consuming. Therefore, how to select the discriminative ones from the feature pool is the key problem. In our paper, we propose to using ranking model to do feature selection.

3 Ranking Model

Age changing could be looked as a time series problem. Obviously, age monotonously increases with time, therefore, the powerful features should be helpful on distinguishing ages at different states within one subject. Inspired by the idea of AGES [8], we also focus on learning the discriminative features within individual subject.

Ranking is widely used in the fields of information retrieval [9] and econometric model [5]. For a ranking problem, it is assumed there is an outcome space $\mathcal{Y} = \{r_1, ..., r_q\}$ with ordered ranks $r_q \succ r_{q-1} \succ ... \succ r_1$, where $\succ$ means larger and denotes the order between different ranks. Generally, one latent continuous function $U(x)$ should be learned to map one sample $x$ into value $r$ in the space $\mathcal{Y}$. In this paper, we propose to use the RankBoost for feature selection, which is based on the ordinal relationship in the pairwise data. We can easily organize the data to get the ordinal information according to the ages on each subject. In this section, we will first introduce how to organize the data, and then we will present the RankBoost on feature selection. Finally the features selected by rank model are used to do age prediction based on different regression models.

3.1 Data Organization

Given a personal age pattern, although it is hard for us to label the exact age on each instance which could be single image or sub-sequence. We can definitely label the ordinal relationship between a pair of instances according to temporal order easily. Without loss of generality, we take the one subject $S_i$ as the example to describe how to organize the data:

Taking account of a subject $S_i$, we get the ordered sequence set $\{Seq_{S_i}\}$ as personal age pattern, and based on $\{Seq_{S_i}\}$, we build pairwise instances $\{(x_k, x_{k+1})\}$ for the ranking model learning to satisfy $R(x_{k+1}) \succ R(x_k)$. Figure 1 shows the example of the age pattern on different subjects $S_i$'s, and the image sequences are organized by the age order. We could extract ordered pairs $\{(x_k, x_{k+1})\}$s which satisfy the age of $x_{k+1} \succ$ the age of $x_k$. The ordered pairs can be exhaustively extracted from one ordinal sequence for ranking model learning. In figure 1, we just show the pairs holding adjacent instances.

The features work well on the ordered pairwise data, they should be discriminative and helpful on the age prediction. Based on our data organizing strategy, we enforce the ranking model to learn the powerful features to rank the aging face images within each subject.
3.2 RankBoost

In this paper, we use the haar-like features to represent facial appearance, so we have thousands of haar-like features. It is untractable to use all the haar-like features in regression model. Moreover, age info is only dominated by parts of facial appearances. Thus, we adopt the RankBoost to build the ranking model over the ordinal pair-wise data. Similar to the boosting learning, the RankBoost [6] aims to select a set of weak rankers to build a strong ranker. Given pairwise sample sets \( \{(x_{i,0}, x_{i,1})\} \), the RankBoost tries to find a ranking function \( H(x) = \sum_t \alpha_t h_t(x) \), where \( h_t(x) \) is a weak ranker, based on the loss function \( \min \sum_{x_0, x_1} \exp(\sum_t \alpha_t (h_t(x_0) - h_t(x_1))) \). The Rankboost additively selects weak rankers by minimizing the exponential loss function, in which the greedy optimization strategy is used to solve it. The detailed algorithm is presented in Algorithm 1.

Algorithm 1 RankBoost Learning procedure

1: Give example image pairs \((x_{i,0}, x_{i,1})\)\(\ldots\)(\((x_{n,0}, x_{n,1})\).
2: Initialize weight \(D_t(i) = 1/N\).
3: for \(t = 0 \ldots T\) do
4: \hspace{1em} Train weak learner using distribution \(D_t(i) = 1/N\).
5: \hspace{1em} Get weak ranking \(h_t : h_t(x) \rightarrow R\), s.t. equation \(\min \sum_{x_0, x_1} \exp(\sum_t \alpha_t (h_t(x_0) - h_t(x_1)))\) = \(\min \sum_{x_0, x_1} \exp(\sum_t \alpha_t (h_t(x_0) - h_t(x_1)))\).
6: \hspace{1em} Choose \(\alpha_t \in R\).
7: \hspace{1em} Update:
8: \hspace{1.5em} \(D_{t+1}(x_{i,0}, x_{i,1}) = \frac{D_t(x_{i,0}, x_{i,1}) \exp(\alpha_t (h_t(x_{i,0}) - h_t(x_{i,1})))}{Z_t}\)
9: \hspace{1em} where \(Z_t\) is a normalization factor.
10: end for
11: Output the final ranking \(H(x) = \sum_t \alpha_t h_t(x)\).

4 Experiment

The FG-NET[1] aging Database contains 1,002 face images from 82 subjects. The ages in FG-NET databases are distributed highly unevenly in wide ranges: 0-69. For each face in the database, we know the corresponding age. For each subject, one image sequence ranked based on the age order is used to extract pairwise data. Rankboost is used to select discriminative features, and different kinds of regression methods are trained based on the selected features. Same as [13], we use 4 folder cross validation to evaluate our method.

The performance of age estimation are usually evaluated by the Mean Absolute Error (MAE). MAE is defined as the mean of the absolute errors between the estimated ages and the ground truth:

\[
MAE = \frac{1}{N} \sum_{i=1}^{N} |\hat{l}_i - l_i|/N
\]

where \(\hat{l}_i\) is the ground truth age for image \(i\), and \(l_i\) is the corresponding estimated age.

1000 haar features are selected by rankboost, Figure 2 shows the relation between training error and number of features, where training error is measured by Relevant Accuracy (RA). The definition of RA is:

\[
RA = \frac{\text{number of correctly ranked relevant pairs}}{\text{number of all relevant pairs}}
\]

To evaluate the performance of the selected haar features, we take Linear Regression, Multi-Layer Perceptron (MLP), and Support Vector Regression (SVR).
Table 1. Performance on the FG-NET database

<table>
<thead>
<tr>
<th></th>
<th>Linear Regression</th>
<th>MLP</th>
<th>SVR</th>
<th>Boosting</th>
<th>Gaussian Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>top 20</td>
<td>7.65</td>
<td>7.49</td>
<td>6.74</td>
<td>6.68</td>
<td>6.87</td>
</tr>
<tr>
<td>top 50</td>
<td>7.16</td>
<td>9.65</td>
<td>5.79</td>
<td>6.68</td>
<td>6.46</td>
</tr>
<tr>
<td>top 100</td>
<td>7.18</td>
<td>12.08</td>
<td>5.68</td>
<td>6.95</td>
<td>5.98</td>
</tr>
<tr>
<td>top 200</td>
<td>8.75</td>
<td>11.40</td>
<td>5.67</td>
<td>6.47</td>
<td>6.07</td>
</tr>
<tr>
<td>top 500</td>
<td>9.41</td>
<td>9.78</td>
<td>5.67</td>
<td>6.66</td>
<td>5.98</td>
</tr>
</tbody>
</table>

Table 2. MAE of different works

<table>
<thead>
<tr>
<th></th>
<th>AGES</th>
<th>RUN</th>
<th>MLP</th>
<th>SVR</th>
<th>Boosting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xin[8]</td>
<td>6.22</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Yan[16]</td>
<td>5.78</td>
<td>10.39</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Our method</td>
<td>-</td>
<td>7.49</td>
<td><strong>5.67</strong></td>
<td>6.47</td>
<td></td>
</tr>
</tbody>
</table>

We also compare with the Additive-Regression (boost). Different size of selected feature set is used on the above regression model. MAE is used to check the performance. SVR gets the best result with MAE 5.67 in our work, the RBF kernel is used with gamma 0.06.

From the result in table 1, we can see the the MAE is very stable while even more features are used. The amazing thing is the when just 50 features are used, the performance of our method is comparable to the results of the state-of-art approaches, 6.22 in Xin’s [8], 5.974 in Suo[13] and 5.78 in Yan’s [16]. In [13], Suo also used four fold cross-validation strategy which is same as ours. For other three works, they use Leave-One-Person-out strategy. Compared with their strategies, four folds cross validation is more challenging.

Table 2 shows more detail experiment results in other work. Especially compared with suo’s[13], our proposed features are simpler and more efficient, 792 features were used in [13], however, only 50 features are used in our proposed method to get better result.

5 Conclusion

In this paper, we propose to using ranking model to select discriminative features based on pairwise data built on personal age patterns. The performance of different regression methods based on the selected features is comparable with the state-of-art works, even better. By proper strategy, simple texture features also get great results.

References