Can Estimate Age Range Using 'a Face a Person'?

Sujing WANG[1], Chunguang ZHOU[1], Mingfang SUN[1], Huiling CHEN[1], Xiaohua LIU[1†], Xujun Peng[2]

[1] *College of Computer Science and Technology, Jilin University, Changchun 130012, China*

[2] *Raytheon BBN technologies, Boston, MA, 02138, USA*

Abstract

After decades of research and development, facial recognition system have attained considerable success in that many real-world applications are implemented, such as human computer interaction and multimedia communication. However, the technique of estimating human age via facial image is still a challenging problem for existing computer vision systems to automatically and effectively predict the exact age of humans. The problem derives from the fact that personalized aging patterns are temporal data and it is difficult to collect sufficient training data. In this paper, we utilize the classical Active Appearance Model and Artificial Neural Network to explore age range estimation using a Face a Person scheme in training phase. The experiments are conducted on the well-known aging database FG-NET. In order to show the effectiveness of the proposed method, the benchmark algorithm AGES, which needs subjects photo temporal sequences as the training set is used to compare against the proposed algorithm. The experimental results show that the performance of our method outperforms that of AGES algorithm.

Key words: Computer Vision; Pattern Recognition; Machine Learning; Neural Network; Age Estimation

1. Introduction

Human faces contain a significant amount of nonverbal information including the identity, emotional state, ethnic origin, gender and age. The ability of using the facial information during interaction makes it possible for humans to recognize and interpret faces and facial gestures accurately in real time. And in the field of computer vision, such as multimedia communication and Human Computer Interaction (HCI), facial attributes also play a key role in real-world applications.

Although lots of efforts have been taken to ensure machines the capability of identifying faces[1], recognizing emotions[2] and genders, etc, very few work have ventured forth to develop an applicable automatic age estimation system, which provides a wide range of uses. For example, in order to restrict internet access and enhance the parental control over websites of adult materials for children, it would be necessary to estimate the age of users and take different actions accordingly. A vending machine could be installed with an age estimation system to recognize underage kids and refuse to sell cigarettes or alcohol[3,4]. In most cases, instead of estimating the exact age of a given face, the age range estimation is enough. For instance, the police are more interested in describing an unknown suspect as "46 to 50 years old" rather than "exactly 48 years old".

No matter age range estimation or exact age estimation, automatic age estimation system from human face images is still a challenging problem which encounters several difficulties[5]:

- **Diversity of aging variation.** Although all people age, they do so in different ways.
- **Dependence on external factors.** Aging variation can be affected by external factors such as health, standard of living, and exposure to extreme weather conditions.

This paper is supported by (1) the National Natural Science Foundation of China under Grant No. 60873146、60973092、60903097, (2) National High Technology Research and Development Program of China under Grant No. 2009AA02Z307, (3) project of science and technology innovation platform of computing and software science (985 engineering), (4) the Key Laboratory for Symbol Computation and Knowledge Engineering of the National Education Ministry of China.

† Corresponding author.

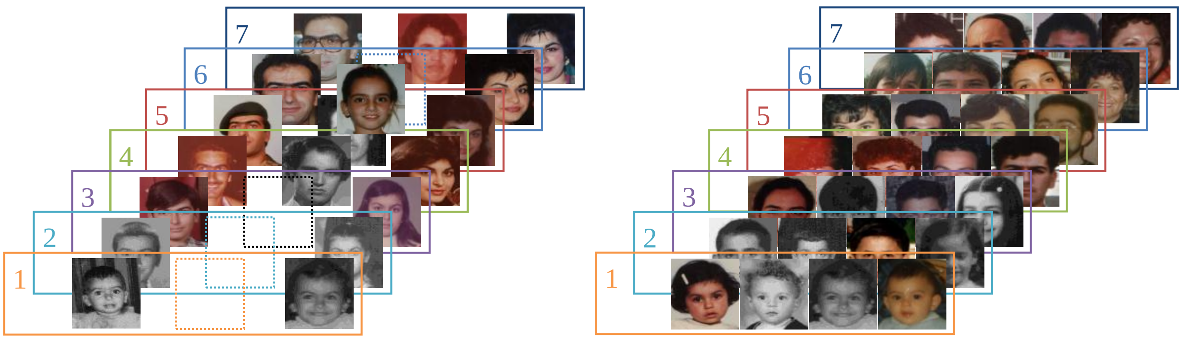
Email addresses: liuxiaohua@yahoo.cn (Xiaohua LIU).

- **Collecting training data.** Because the process of aging is slow, the collection of suitable data set for training and recognizing this type of variation is difficult.

Fortunately, FG-NET Aging Database[6] relieves the difficulty of collecting training data in some degree.

With the countless areas for which the age estimation systems can be used, there is only limited number of literatures cover the automatic age estimation. Existing methods for age estimation using face image can be roughly divided into 4 categories: anthropometric model[7, 8], aging pattern subspace[9-12], age regression[4, 13] and age classification[4, 14]. The anthropometric model uses the cranio-facial development theory and facial skin wrinkle analysis. The growth related changes of face shape and texture patterns are measured to categorize a face into several age groups. An aging pattern is a sequence of personal face images sorted in time order. The aging pattern subspace method models a sequence of personal aging face images by learning a subspace representation, to deal with incomplete data set such as missing ages in the sequence. The age of a subject's facial image is determined by the projection onto the subspace (linear, non-linear kernel or multi-linear) that can best reconstruct the subject's facial image. However, the problem lies in the fact that it is rather difficult to collect sufficient sequences of a single individual's aging face images. To the regression method, facial features are extracted from an appearance-based model.

Existing age estimation algorithms require all face images taken at diverse age grades for every training sample, which is essentially their inevitable difficulty. As shown in Figure 1, there are 7 different age ranges. The second person has lots of missing images in various age ranges in 'Multi Face A Person' database. The data set's characteristics are: (1) high missing data; (2) small size samples. This characteristics lead to the difficulties of age estimation.



(A) Multi-Face A Person (B) A Face A Person

Figure 1 (A) Multi-Face A Person and (B) A Face A Person

In order to estimate age, Ben et al.[15] used a single face a person in training set. But they did not compare the performance with the others methods.

2. Proposed method

3.1 Feature extraction

Prior to age estimation, we use Active Appearance Model (AAM), which combines the shape and intensity information, to extract features for a given facial image. The shape model is trained on 68 manually labeled key points on each face image as shown in Figure . To the intensity model, each face image is aligned to the mean face shape (i.e. the mean of the key points) of the training set. It has been shown that the Appearance Model is robust against many facial variations such as illumination, view angle, and facial hair. More details of AAM can be found in[16].

Principal Component Analysis (PCA) on the deviations of each example from the mean example is used to reduce the size of features. Then, training examples can be reconstructed/parameterized using Equation (1)

$$X = X_m + Pb \qquad (1)$$

where X is a vector which represents the shape or intensity pattern of a training example, X_m is the mean of training samples, P is the matrix of eigenvectors, and b is a vector of weights, or model parameters. Edwards et al.[16] described in detail how this type of model can be used for modeling through combination of shape and intensity variation of face images. As an extension to the basic technique, Edwards et al.[17] described how color models can be generated by using the intensity pattern of the RGB component of each pixel in the facial region. In this paper, the final dimension of feature vector for each facial image is 181, which includes 27 shape-related features and 154 texture-related features.

Figure 4. 68 key points on face image

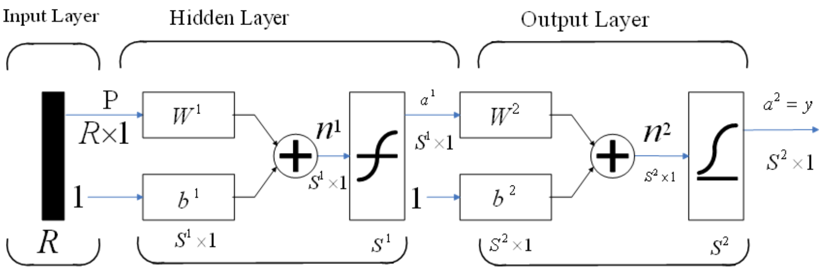


Figure 5. The Structure of Neural Network

3.2 The structure of neural network

Neural network is used in our system to estimate the age of face image. A typical multi-layer perceptron (MLP) neural network has an input layer, one or more hidden layers, and an output layer. The number of neurons in the input layer is determined by the number of inputs (features) of the network, while the number of outputs determines the number of neurons in the output layer. The number of neurons in the hidden layers can be determined by trial and error, network growing rates, or network pruning.

In this paper, we use a neural network whose structure is shown in Figure 5. The input vector P is a vector with length R, where $R = 181$ is determined by the size of features as described in Section 1. There are S^1 neurons in hidden layer. The hidden layer includes a $S^1 \times R$ weight matrix W, a bias vector b with length S^1, a summator, a transfer function Hyperbolic Tangent Sigmoid and the output vector a^1 with length S^1. The vector a^1 is the input vector of the output layer. The output layer structure is similar to the hidden layer structure, except for the transfer function which is Log-Sigmoid. The output of the neural network is the vector a^2. The length of a^2 is determined by the number of the age ranges. The batch mode is used to train the network. In batch mode the weight and biases of the network are updated only after the entire training set has been applied to the network. The gradients calculated for each training example are added together to determine the change of the weights and biases. The batch algorithm traingdm (Batch Gradient Descent with Momentum) is used in our experiment. The algorithm for feedforward networks often provides faster convergence. Momentum allows a network to respond not only to the local gradient, but also to recent trends in the error surface. The performance function used for training the neural network is the mean sum of squares of the network errors.

3. Experiments

The experiments are implemented through the Leave-One-Person-Out (LOPO) mode, i.e. in each fold, the images of one person are used as the test set and those of the others are used as the training set(about $(1002/82)\times 81 \approx 990$ images). After 82 folds, each subject has been used as test set once, and the final results are calculated based on all estimations and the network error rate is about 0.04. Thus, the system is tested in the similar manner as the real application, i.e. the subject for whom the algorithms attempt to estimate his/her age is previously unseen in the training set. Meanwhile, the relatively limited data set can be adequately utilized. The performance of age estimation is measured by two different metrics [5] : the Mean Absolute Error (MAE) and the Hit Rate. The MAE is defined as the average of the absolute errors between the estimated ages and the ground truth ages, $MAE = \sum_{k=1}^{N} |\hat{l}_k - l_k| / N$, where l_k is the ground truth age for the test image k, $\hat{l}_k$ is the estimated age, and N is the total number of test images. The Hit Rate is the percentage of the exactly correct estimations.

Two separate experiments are conducted in order to investigate the relationship between the size of age range and the performance of the proposed method. In the first experiment, the age range from 0 to 49 is divided into 5 groups, each of which ranges 10 ages, i.e. 0-9, 10-19, 20-29, 30-39 and 40-49. In the second experiment, the age ranges from 0 to 19 is divided into 5 groups, each of which contains 4 ages, i.e. 0-3, 4-7, 8-11, 12-15 and 16-19. With the reduction in the number of samples for each age range, the recognition rate declines, as shown in Figure 2 to Figure 37. In the second experiment, the numbers of samples are almost equal. Thus, the recognition rates for each age range are almost equal, except Age Range 3 and Age Range 4. The faces of children whose ages are between 8-14 years are at their puberty and have relatively large changes in age estimation. This is one of the reasons why the recognition rate for Age Range 3 and Age Range 4 are relatively lower. The MAE of the first experiment and the second experiment are 0.4923 and 0.6437, respectively. And the Hit Rates of them are 45.02% and 51.35%, respectively. The MAE of the second experiment is higher than the MAE of the first experiment. The reason is that the age range in the second experiment is narrower than the age range in the first experiment.

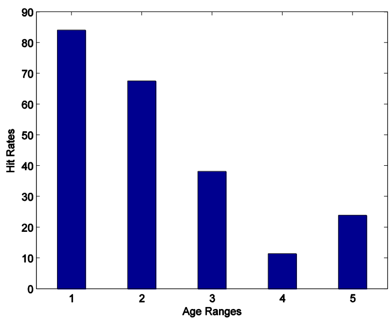


Figure 2. Hit rate for each age range, in the first experiment

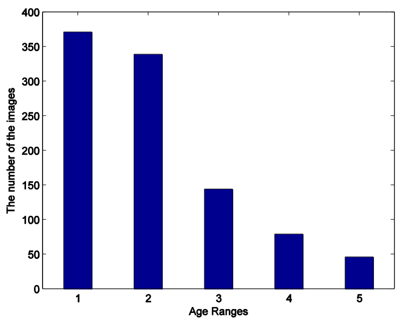


Figure 3. The number of the images in each age range, in the first experiment

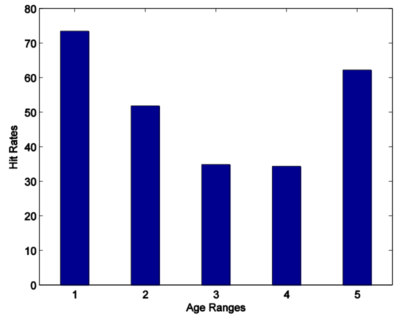


Figure 4. Hit rate for each age range, in the second experiment

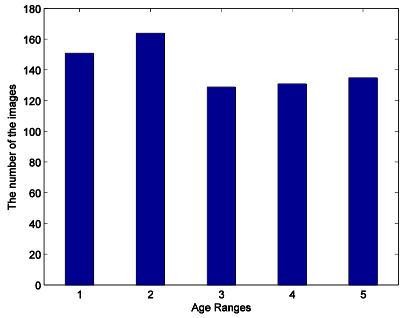


Figure 5. The number of the images in each age range, in the second experiment

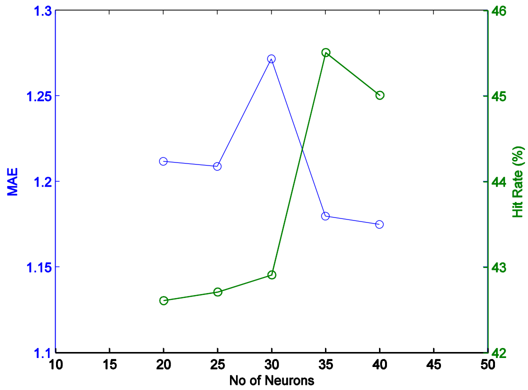


Figure 6 Relationship between the number of neurons in the hidden layer and the performance

In order to see the effects of the structure of Neural network on the performance of age estimation, the ages (0-69) are divided into 14 ranges where each of them contains 5 continuous age, i.e. 0-4, 5-9, ..., 65-69, and the Neural network, which has 181 neurons ($R = 181$) in the input layer and 14 neurons ($S^2 = 14$) in output layer corresponding to 14 age ranges, is used in our experiments for testing. Figure 6 shows the relationship between the number of the neurons in the hidden layer and the network performances (MAE and Hit Rate). It can be seen that the hit rate reaches its peak and the MAE drops to the second lowest value, which is very close to the lowest point, when the number of the neurons in hidden layer is 35.

The performances of the proposed algorithm and the algorithm AGES[11] in age range estimation are compared in Table 1. It can be seen that the proposed method achieves the higher hit rate and better MAE than algorithm AGES.

Table 1. Performance of the CEO and others in Age Range Estimation

Method	MAE	Hit Rate(%)
CEO	1.17	45.51
AGES	1.79	31.53

错误!未找到引用源。 shows that the age ranges estimated by CEO with 35 units in the hidden layer versus the actual age ranges. It can be seen that the CEO predictions give larger age values for some people in 1 age ranges, smaller age values for almost all of people in 7-14 age ranges. It can also be seen that with the decreasing number of training samples in certain age ranges, the recognition rate is lower in corresponding age range. The number of samples, the number of correct recognition, the hit rate, and MAE in every age ranges are tabulated in Table 2. Figure 9 and Figure 10 show the accuracy and the number of the images in each age ranges, respectively. It can be also seen from Figure 9 and Figure 10 that with the reduction in the number of samples for each age range, the recognition rate declines. Comparing with Figure 4 and Figure 9, it shows that with the reduction in the number of samples for each age range, the recognition rate declines. The experimental results also support the fact that the faces of children between 8-14 years old undergo relatively large changes of accuracy because of puberty.

Table 2. The number of samples, the number of correct recognition, hit rate, and MAE in every age ranges

Age Range	1	2	3	4	5	6	7
Number of Samples	193	178	174	165	84	60	46
Number of Correct Recognition	172	92	56	108	14	9	1
Hit Rate(%)	89.11	51.68	32.18	65.45	16.66	15	2.17
MAE	0.12	0.56	0.77	0.47	1.13	1.93	2.95
Age Range	8	9	10	11	12	13	14
Number of Samples	33	28	18	12	3	6	2
Number of Correct Recognition	2	1	0	1	0	0	0
Hit Rate(%)	6.06	3.57	0	8.33	0	0	0
MAE	3.60	4.35	4.94	6.83	6.33	7.16	11

From Table 2 we can see that the hit rates are zeros in Age Range 10, 12, 13 and 14, but the MAEs are increasing as the increasing of ages. Is there a certain relationship between MAE and age range? Is there a certain relationship between MAE and the number of training samples in age range? In order to find the answers, two experiments are conducted in the age range from 0 to 39. The ages (0-39) are divided into 8 groups and each of which contains 5 continuous age, i.e. 0-4, 5-9, ..., 35-39. We collect 33 samples for each age range in the first experiment, denoted as Experiment A, and collect all available samples for each subject in the second experiment, denoted as Experiment B. It can be seen from Figure 8 that the MAE of Experiment A is increasing steadily as the age increases and the MAE of Experiment B is dramatically changing. This shows that the MAE is more relied on the number of the samples rather than the age range.

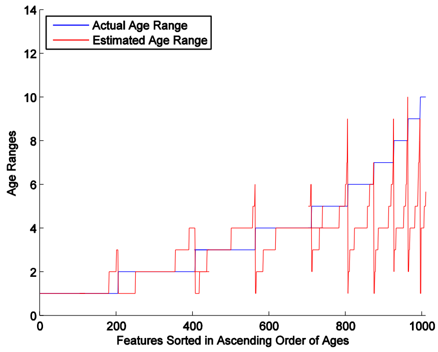


Figure 7 Age Ranges Predicted by CEO versus the Actual Age Range

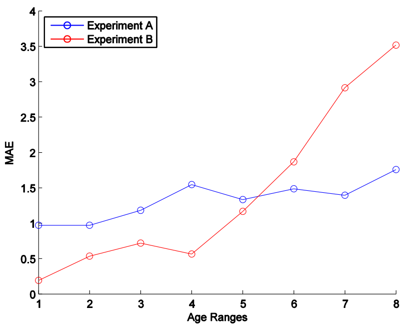


Figure 8 The MAEs for each age range in Experiment A and Experiment B

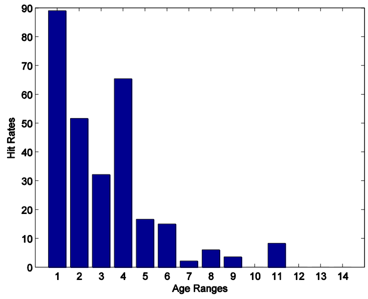


Figure 9 Hit rate for each age range

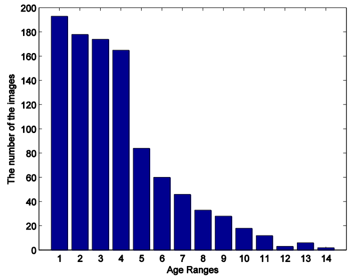


Figure 10 The number of the images in each age range

4. Conclusion

Automatic human age estimation has numerous potential applications in human computer interaction and multimedia communication. However, the age estimation problem is challenging; the collection of sufficient training data, many sequences of personal aging face images in chronological order, for age estimation is extremely laborious and difficult. In this paper, we utilize the classical Active Appearance Model and Artificial Neural Network to explore the problem – ‘Can estimate age range using a Face a Person in training set?’. The experimental results show that the proposed method achieves higher hit rate and better MAE than algorithm AGES.

References

[1] Lanitis A, Taylor C J, Cootes T F. Automatic interpretation and coding of face images using flexible models[J]. 1997: 743-756.

[2] Pantic M, Member S, Rothkrantz L J M. Automatic Analysis of Facial Expressions: The State of the Art[J]. 2000: 1424-1445.

[3] Yang M, Moghaddam B, Moghaddam B. Gender Classication with Support Vector Machine[Z]. 2000306-311.

[4] Lanitis A, Draganova C, Christodoulou C. Comparing different classifiers for automatic age estimation[J]. Ieee Transactions on Systems Man and Cybernetics Part B-Cybernetics. 2004, 34(1): 621-628.

[5] Lanitis A, Taylor C J, Cootes T F. Toward automatic simulation of aging effects on face images[J]. Ieee Transactions on Pattern Analysis and Machine Intelligence. 2002, 24(4): 442-455.

[6] FG-NET AGING DATABASE. http://www.fgnet.rsunit.com/[Z]. : 2009.

[7] Kwon Y H, Lobo N D. Age classification from facial images[J]. Computer Vision and Image Understanding. 1999, 74(1): 1-21.

[8] Ramanathan N. Modeling age progression in young faces[Z]. 2006387-394.

[9] Geng X, Zhou Z H, Smith-Miles K. Automatic age estimation based on facial aging patterns[J]. Ieee Transactions on Pattern Analysis and Machine Intelligence. 2007, 29(12): 2234-2240.

[10] Geng X, Smith-Miles K, Zhou Z. Facial age estimation by nonlinear aging pattern subspace [Z]. Vancouver, British Columbia, Canada : ACM, 2008721-724.

[11] Geng X, Zhou Z, Zhang Y, et al. Learning from facial aging patterns for automatic age estimation [Z]. Santa Barbara, CA, USA : ACM, 2006307-316.

[12] Geng X, Smith-Miles K. FACIAL AGE ESTIMATION BY MULTILINEAR SUBSPACE ANALYSIS[Z]. NEW YORK: 2009865-868.

[13] Guo G D, Fu Y, Huang T S, et al. Locally adjusted robust regression for human age estimation[J]. 2008 Ieee Workshop on Applications of Computer Vision. 2008: 112-117.

[14] Guodong G, Guowang M, Yun F, et al. Human age estimation using bio-inspired features[Z]. 2009112-119.

[15] Ben S L, Su G D, Wu Y S. A Two-Step Selective Region Ensemble for Facial Age Estimation[J]. ADVANCED INTELLIGENT COMPUTING THEORIES AND APPLICATIONS, PROCEEDINGS. 2008, 5227: 622-632.

[16] Edwards G J, Lanitis A, Taylor C J, et al. Statistical models of face images - improving specificity[J]. Image and Vision Computing. 1998, 16(3): 203-211.

[17] Edwards G J, Cootes T F, Taylor C J. Advances in Active Appearance Models[Z]. Kerkyra, Greece: IEEE, 1999: 1, 137-142.