Design Sparse Features for Age Estimation using Hierarchical Face Model

Jinli Suo[1,3] Tianfu Wu[3], Songchun Zhu[3] Shiguang Shan[2] Xilin Chen[2] Wen Gao[4,1]

jlsuo.lhi,tfwu.lhi,sczhu@gmail.com sgshan,xlchen@ict.ac.cn wgao@pku.edu.cn

[1]Graduate University of Chinese Academy of Sciences(CAS), 100190, China
[2]Key Lab of Intelligent Information Processing of CAS,
Institute of Computing Technology, CAS, Beijing, 100190, China
[3]Lotus Hill Institute for Computer Vision and Information Science, 436000, China.
[4]Institute of Digital Media, Peking University, 100871, China

Abstract

A key point in automatic age estimation is to design feature set essential to age perception. To achieve this goal, this paper builds up a hierarchical graphical face model for faces appearing at low, middle and high resolution respectively. Along the hierarchy, a face image is decomposed into detailed parts from coarse to fine. Then four types of features are extracted from this graph representation guided by the priors of aging process embedded in the graphical model: topology, geometry, photometry and configuration. On age estimation, this paper follows the popular regression formulation for mapping feature vectors to its age label. The effectiveness of the presented feature set is justified by testing results on two datasets using different kinds of regression methods. The experimental results in this paper show that designing feature set for age estimation under the guidance of hierarchical face model is a promising method and a flexible framework as well.

1. Introduction

Although many aspects of face variations, such as identity, pose, et al., have been widely and extensively studied in the literature of computer vision and pattern recognition, age has not yet been well explored. Automatic age estimation is attracting increasing research interests recently, mainly because its wide potential applications, including developing group specific human computer interfaces for human-machine system, improving recognition system efficiency on large databases, performing age specific customer analysis, et al. This paper presents a flexible methodology to design features suitable for automatic age estimation by building a three-level hierarchical face model.

1.1. Previous works

In the literature, automatic age estimation has not been widely studied compared with other variations of human faces. Preliminary works[1, 2, 3, 4] formulate age estimation as a classification problem and obtain coarse estimation by using geometric features to discriminate adults from children and texture features in skin zones to discriminate young adults from senior ones. Later researchers attempt to conduct more precise age estimation. Among them, active appearance model(AAM)[5] is the most popular face model used to project face images onto a low dimension parametric space and perform estimation on face parameters[6]. Geng et al. [7] define 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 proper pattern. Similarly, Fu[8] learns low dimensional aging manifold from image intensities and age estimation is formulated as a regression problem in manifold space. Some other researchers adopt traditional discriminative methods[9], using image intensities directly or other features exhaustively extracted from images to do estimation. In order to advance research on face aging and evaluate performance of age estimation methods, some researchers made large efforts on building aging face database[10, 11].

In the existing works, although much progress has been made, there are some weaknesses in two-folds: (i)The previously adopted global face models in terms of a single level AAM are usually not expressive enough to represent details critical to age perception, and as an important cue for age perception, hair is often excluded. (ii)Few efforts have gone into the feature design problem which is essential to age estimation. Often, besides the AAM, image intensities or dense features exhaustively extracted from image patch are used, leading to high dimensional feature vectors and causing the training process time consuming and demand large database to overcome over-fitting.

1.2. Motivation and proposed approach

The aforementioned problems motivate us to adopt the hierarchical face model and design sparse feature set based

978-1-4244-2154-1/08/$25.00 2008 IE

on it. The hierarchical face model is first proposed in Xu[12]. In this paper, we build up the hierarchical face model by extending the global AAM model with two additional levels: (i) The first level is a global AAM model for face images at low resolution, the same as representation traditionally used in age estimation. (ii) At second level, we first cluster each facial component into different types and add a separate AAM model for each type as a refined description of facial components. This accounts for middle resolution. (iii) Further, in order to describe facial details perceptible at high resolution, skin zones are further refined with some details(e.g.wrinkles and blobs). In addition, according to the importance of hair to age perception[13], we also include hair area in the model. These extensions effectively describe the large variations of facial components and skin zones across large age ranges.

Base on the presented hierarchical graphical face model, we design a set of features for age estimation for images at different resolutions. The details of feature extraction is described in Sec.4. For the feature extraction occurs only on the parts with large variations across aging process, our features are of much lower dimension than image intensities or dense features exhaustively pursued by traditional discriminative methods. The low dimensional features alleviate the problems of long training process and demanding store memory in learning from high dimensional features. Experimental results show that the designed features have strong discriminative power and give promising performance as is shown in Fig.5.

In general when the training set is small, high dimensional features will suffer from "curse of dimension" and heavily deteriorate the performance of estimation methods. In contrast the designed features exclude disturbance of a large amount of unrelated features and alleviate over-fitting to some extent(see results shown in Table.4).

2. Method overview

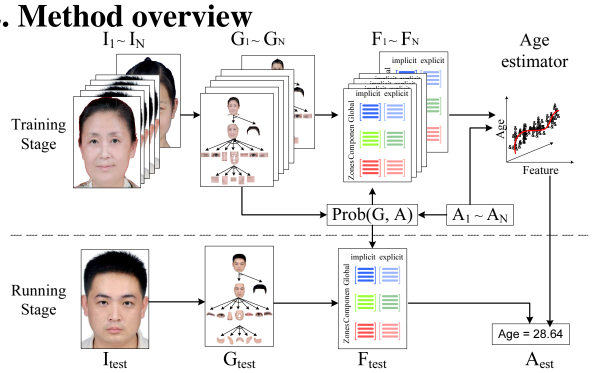


Figure 1. Flow chart of our framework.

The flow chart of our framework is as shown in Fig.1: In training stage, we collect a set of face images I_1 to I_N as training set and N denotes the number of training samples. Firstly graph representations G_1 to G_N of I_1 to I_N are computed. From these graphs, a joint probability of the graph representations and age labels can be learned as $Prob(G, A)$, under whose guidance we extract feature vectors F_1 to F_N from G_1 to G_N for age estimation. Finally an age estimator is trained from the feature vectors and their correspondent age labels A_1 to A_N. In running stage, graph computation and feature selection are the same as in training stage, and supplied to the estimator to compute the estimated age, which takes the form of a nonnegative real number. We give a general explanation for the key modules:

(1)*Computation of graph representation.* In our estimation task, we adopt hierarchical graph model for face representation which is composed by a set of nodes and edges connecting them. This graph model decomposes face into semantically meaningful parts represented as graph nodes at three levels, which describe face details at different resolutions. There exists vertical edges between nodes at adjacent levels and horizontal ones between nodes at the same level describing the constraints and spatial relationships. In this graph representation, a face instance is presented by the parameters of nodes and constraints between them.

(2)*Learning statistics of graph parameters.* From parameters of G_1 to G_N and age labels A_1 to A_N, we can learn a joint probability $\text{Prob}(G, A)$, which provides guidance to feature selection for age estimation. We extract features from the parameters with large variations along age ranges.

(3)*Feature design guided by hierarchical model.* Inspired by the observation that humans usually perform age estimation from local and informative features, including general face and skin attributes, wrinkles, ratios between metrics from facial landmarks, et al., we extract features from parameters of graph nodes at different levels(illustrated with different color cues in Fig.1). Besides parameters as explicit features we also compute some implicit ones, we differentiate them in Fig.1 with bars of different saturations. For feature extraction is performed under guidance of face model, we call it feature design.

(4)*Age estimation in feature space.* Considering the temporal and continuous property of aging pattern, we formulate the problem of age estimation from designed features as a regression problem. For aging process is continuous, the estimation result takes the form of nonnegative real number.

3. Hierarchical model guides feature design

In this section, we first describe the adopted hierarchical face model and its guidance on feature design, after which we give the formulation of age estimation problem.

3.1. Hierarchical face model

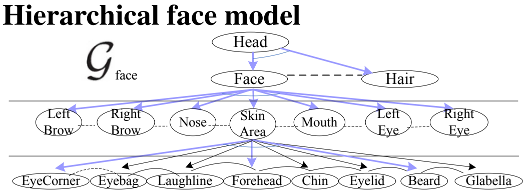


Figure 2. Hierarchical face model.

In our framework, we represent all human faces with a hierarchical graph model $\mathcal{G}_{\text{face}}$, as shown in Fig.2. The model represents human faces with global appearance at first level, variety of facial components at second level, and detailed information in eight skin zones at third level. By selecting a specific topology and assigning proper attributes to nodes and edges of $\mathcal{G}_{\text{face}}$, it turns into a parse graph G(e.g. graph composed by blue edges and nodes connected with them in Fig.2), which generates a face instance I.

$$G \longrightarrow I \tag{1}$$

Adopting coarse-to-fine idea, graph G is controlled by three hidden variables, each for generation of information at one resolution.

$$G = \{w_1, w_2, w_3\} \tag{2}$$

Here $w_1 = \{w_{\text{hair}}, w_{\text{face}}\}$ describes the general appearance of human face and hair. w_2 refines the appearance of facial components(eye, nose, et al.) and w_3 further refines the attributes of different facial zones, including wrinkles, pigments, marks, et al.

At each level, four kinds of attributes are described in our model: topology, geometry, photometry, and configuration. Correspondent parameters are separately denoted as T_i^{top}, T_i^{geo}, T_i^{pht} and T_i^{cfg}.

$$w_i = \{T_i^{\text{top}}, T_i^{\text{geo}}, T_i^{\text{pht}}, T_i^{\text{cfg}}\} \tag{3}$$

The aforementioned parameters form a unified parametric representation of faces across all ages, the features for age recognition are extracted from these parameters.

The prior probability and likelihood of the face model is as follows:

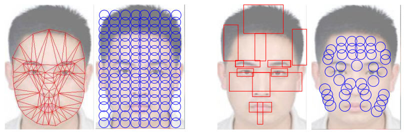
$$p\{G;\Theta\} = p(w_1;\Theta_1)p(w_2|w_1;\Theta_2)p(w_3|w_2;\Theta_3) \tag{4}$$

$$p\{I|G\} = \prod_{i=1}^{3} p(I_i|w_i) \tag{5}$$

Here Θ is the parameter constraining the relationships between nodes at adjacent levels and at the same level.

3.2. Feature selection under guidance of face model

Hierarchical face model is a compositional model, which decomposes face into parts according to prior knowledge. Both the nodes and the relationships between them have semantic meaning and we can design a candidate feature set heuristically for age classification. As is illustrated in Fig.3(a) discriminative approaches extract features exhaustively by applying different filters to pixels sampled densely from the lattice. This strategy results in high dimension feature vector, which causes demanding store memory and long training process. In contrast, our approach designs appropriate features at selected positions, both the features and the positions are from the guidance of prior knowledge. As in Fig.3(b), the gradients(blue circles) are extracted in the wrinkle zones, geometric metrics(red rectangles) are extracted as features at facial components, hair, et al. The designed features consistent with prior knowledge of age perception and has dimension much lower than features from discriminative methods.



(a) Exhaustive Pursuing (b) Guided Design

Figure 3. Comparison between dense features pursed exhaustively and sparse features guided by hierarchical face model.

Some parameters are informative for age classification, we add them directly into the candidate feature set, this type of features are called *explicit features*. Also we can apply some filters onto these parameters and generate a series of discriminative features, we call them implicit features.

3.3. Formulation

For a given image I, our objective is to estimate its age A, which is inferred from a regression model.

$$A = \mathcal{R}[F(I)] \in [0, 100] \tag{6}$$

Here R is the regression function and $F(I)$ is the feature set extracted from I.

$$F(I) = \{(M_i^k, \omega_i); k = 1, 2, ..., D \quad i = 1, 2, 3\} \tag{7}$$

Here i indexes graph level and M denotes filters operated on hidden variable ω, which is the hidden variable for graph representation of I, as in Sec.3. D is the total number of feature numbers. The details of feature candidate design are displayed in the following section.

Hidden variable ω is computed by maximum posterior probability:

$$\omega^* = \arg\max p(\omega|I) \tag{8}$$

4. Implementation aspects

In this section we give a simple review of graph parameters and explain the procedure of feature design. Guided by the hierarchical model, we apply specific filters to different parameters instead of exhaustively pursuing all filters densely over the image lattice. In our approach, we extract four type of features:

Topological features. At three levels from coarse to fine, topological variables represent the index of hair styles, cluster belonging of each facial component and existence of wrinkles of pigment and marks. These parameter are informative features for age perception. For the topological parameters are described by enumerate variables and no partial order should exists among them, we encode these enumerate variables with bi-value vectors as exemplified in Fig.4(b) to ensure distances between them to be equal.

Geometric features. The global AAM at first level and mixed AAMs of facial components at second level give precise localization of facial landmarks, which describe the geometry of human faces. Hair shapes and wrinkle curves are also described by a set of landmarks, as shown in Fig.4(c). With age increasing, geometric changes occur to both facial contour and components: the geometric changes in formative years is especially apparent, while for adult there is only subtle changese.g.drops of eye-corner and mouth-corner, moving up of hairline, growth of wrinkle length et al). Besides the geometric parameters as explicit features, we apply some filters to compute some implicit features, the filters include area, angles, size of out-rectangle, et al.

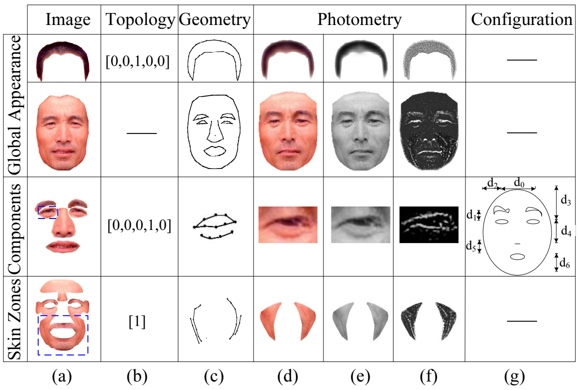


Figure 4. Four types of features at three levels for age estimation.

Photometric features. Photometric parameters describe rich information of face appearance and provide important information to age perception. We compute three cur images from each face image: color, low-frequency intensities and high frequency intensities, as shown in Fig.4(d,e,f). Instead of exhaustively pursuing dense features, we purse different features for nodes at different levels.

For first level, we apply some filters to extract some global statistics about skin and hair attributes from color and low frequency intensities. At second level for refined facial components description, the gradients of facial components provide useful information of component aging. Photometric variables at the third level accounting for the details at different skin zones are important descriptors for age estimation of adult face. We compute statistics from three cue images in each zone.

Configural features. Our face model has a set of distance metrics as configural parameters constraining the spatial relationships among facial components at second level. As is described in [14], Anthropometry statistics provide strong evidence for age estimation of children. We compute a set of ratios from between-component distances as configural features(see Fig.4(g)). These metrics have been used in previous work[1, 2, 3, 4] and proved to be effective.

5. Experiment results

In this section we first give a brief introduction to two databases used in this paper and measures for performance evaluation. After that we design a human experiments as an intuitive support to the idea of estimating age on hierarchical face model and conduct a series of experiments to evaluate our estimation performances in different cases.

5.1. Dataset and quantitative evaluation

We collect a dataset containing 8,000 color face images of Asians, with equal number of males and females and one image per subject. This dataset covers a large age range from 9 to 89 and contains roughly 100 images for each age. Faces in this dataset is at resolution about 120 $\times$ 160 pixels and with little illumination and pose variations. Another dataset is FGNET[10], which is one of the most well-known publicly available face aging databases, composed of 1,002 face images of 82 subjects, whose ages range from 0 to 69.

For each face in the datasets we know its actual age, 90 face landmarks are labeled automatically and rectified by labelers(One can refer to Liang's[15] work for a precise automatic face landmark localization algorithm). Based on the landmarks, we develop an automatic parsing program and extract 792 features for age estimation. On our database, we perform five folds of experiments, each time we randomly select half face images as training samples and the rest for test, the averaged estimation result is computed as the final results. On FGNET, we perform four folds cross validation for performance evaluation.

The performance of age estimation are usually evaluated by two different measures: the **Mean Absolute Error(MAE)** and the **Cumulative Score(CS)**. MAE is defined as the mean of the absolute errors between the estimated ages and the ground truth: $MAE = \sum |l_k - l_k^*|/N$, where l_k^* is the ground truth age for the test image k, l_k is its estimated age and N is the total number of test images. The cumulative score is defined as $CS(j) = N_{e\leq j}/N * 100\%$, where $N_{e\leq j}$ is the number of test images on which the age estimation makes an absolute error no higher than j years. We propose that estimation error over 15 years is unacceptable, and plot CS curves with $j \leq 15$.

5.2. Relative contributions of different parts

We conduct a human experiment to learn the relative contributions of facial parts to age perception of whole face. In this experiment, we randomly select 500 frontal face images from our dataset and divide faces into parts according to graph representation. Ten subjects are presented with the images and asked to perform age estimation to the whole face images and each part separately. A statistical learning methods—Multi Regression Analysis(MRA) is performed on the estimation results to study the relative contribution of each part, analysis result is shown in Table.1.

Here R^2 is a metric measuring the validness of the model. Large value of $R^2 = 0.907$ indicates that our model includes most of the cues related to age perception. The large contributions of skin zones and hair features validate the correctness of extending AAM model with details at high resolution and hair features. Also the relative contributions of each part, provide some cues of face age perception and give some guidances for feature design.

Table 1. Relative contribution of facial parts to age perception of whole face age

	face	eye	configuration	hair	forehead	laughline	eyecorner	brow	eyebag	mouth	eyelid	nose
β	0.374	0.218	0.179	0.113	0.087	0.083	0.058	0.041	0.013	0.009	0.002	0.0012

5.3. Performances of different regressions

	Human	ALR	MLP	SVR	Boosting
Our DB	4.153	4.7413	4.6754	6.4505	4.9603
FGNET	5.65	8.02	5.974	6.97	7.89

Table 2. MAE of different classifiers,

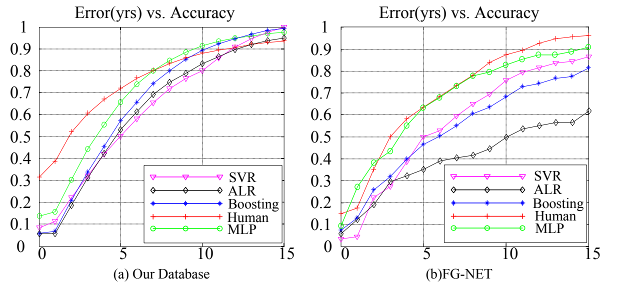


Figure 5. Cumulative scores of four types of estimation methods.

To validate the effectiveness of the designed features, we conduct a comparative study on performances of four kinds of regressions on the designed feature set: Age group specific Linear Regression(ALR), Multi-Layer Perceptron(MLP), Support Vector Regression(SVR) and Logistic Regression(multi-class Adaboost[16]). Human estimation result is also included for comparison. The MAE and CS plot are separately displayed in Table.2 and Fig.5.

From MAE and CS curve of human beings(averaged results of 10 subjects) on our own database, one can see that there exists apparent ambiguity in age perception. Human estimation has an error about 4-5 years, which can be taken as a baseline of automatic age estimation. Among these regressions, MLP gives the highest performance with MAE of 4.68 years, and with error tolerance of 10 ages estimation rate attains 91.6%. The other regressions give results a little lower. In all, these approaches give satisfactory performances in this aging dataset collected under controlled condition, which are a little worse than human estimation.

We also perform the same experiments on FGNET to measure the generalization ability of our estimation approach. For FGNET database has some characteristics different to our database, we adopt some strategies for adaption: (i)we exclude color features to tolerate the gray scale images and discard hair features for the difficulty of extracting hair area from complex background. (ii)There are large external variations(e.g. pose, illumination, expression variations, occlusions, et al.) in the face images, we use RANSAC(Random Sample Consensus) strategy during training to overcome over fitting caused by outliers(with large external variations or of low quality). (iii)There exists large differences on number between images of different ages. the samples of subjects over 40 years is much less than of children and young adults. We reweight the samples at different ages by sampling during training process.

The experiment results are shown in Fig.5(b). Compared to results on our own database, the result is a little worse, which may be caused by following reasons: (1)Some images in FGNET dataset is of low resolution and informative features are not extracted effectively. (2)Samples of some ages don't have enough data for training of regression model, linear regression and boosting suffer from lacking enough training data. (3)Large pose and illumination variations and expression exist in some images.

We plot the CS curves of our best result(MLP) and three existed results: WAS(Weighted Appearance Specific)[6], AGES(AGing pattErn Spaces)[7] and RUN(Nonlinear Regression with and Uncertain Nonnegative Labels)[9]. From Fig.6, one can find that MLP gives MAE of 5.974 years, which is comparable to the results of the-state-of-art approaches, 5.78 years in Geng's[7] and 6.22 years in Yan's[9]. Especially, our results is obtained by four-folds cross validation, which is more challenging then Leave-One-Person-Out strategy used in other three works.

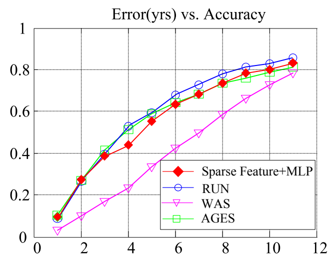


Figure 6. Cumulative scores for different algorithms.

5.4. Estimation performances at three resolutions

In this experiment, we conduct three experiments to study the performances of our age estimation approach on images at different resolutions. We denote our dataset with 8,000 images at resolution 120 pixels $\times$ 140 pixels as set A. By down-sampling we obtain two datasets at decreasing resolutions: set B at 60 pixels $\times$ 70 pixels and set C at 40 pixels $\times$ 47 pixels. We test performance of four types of regression approaches on datasets A, B anc C separately. MAE on different age groups is as shown in Table.3, in which row L, M and H separately represents low, high and middle resolutions and the result with large improvement than lower resolution is displayed with boldface.

From the estimation results on face images at different resolutions, we see that mean absolute error decreases with increasing image resolutions, in other words, estimation accuracy improves with features at higher resolution added. The results show that middle resolution features(attributes of facial components) reduce MAE largely at two age groups: 30-40 and $\geq$ 60, and high resolution features(mostly wrinkles and pigments, marks, et al.) bring significant improvement at three age groups: $\leq$ 20 30-40 and $\geq$ 60. These large improvements may indicate the importance of correspondent features at these groups.

		Age Groups					
		≤ 20	20-30	30-40	40-50	50-60	≥ 60
ALR	L	4.24	3.61	6.16	4.64	4.405	6.794
	M	4.33	3.56	**5.66**	4.25	4.36	**6.47**
	H	**3.96**	3.73	**5.06**	4.22	4.17	**6.07**
MLP	L	4.57	3.79	7.02	4.79	5.88	5.96
	M	4.71	3.84	**6.22**	4.61	5.98	**5.37**
	H	**3.57**	**3.39**	**5.40**	4.54	5.89	**4.74**
SVR	L	5.64	3.27	6.73	12.01	5.76	7.97
	M	5.49	3.16	**5.917**	10.87	5.79	**7.84**
	H	**5.08**	2.93	**5.39**	10.71	5.80	**6.95**
Boost	L	3.85	3.10	5.36	4.62	5.21	7.5
	M	3.79	2.95	**4.83**	4.78	4.85	**6.47**
	H	3.60	2.91	**4.28**	**4.33**	4.78	**6.14**

Table 3. MAE of four estimation methods at three resolutions.

5.5. Performances on training sets of different sizes

In this experiment, we test the performances of age estimators learned from training set of different sizes. By randomly selecting a subset from our dataset(keep the proportion of each age unchanged) as training set and the rest for test, a comparative study of the performances is conducted. In four experiments, number of dataset is separately 4,000, 2,000, 1,000, 500 and four aforementioned regression methods are tested, experiment result is shown in Table.4

	Size of Training Set			
Approach	$4,000$	$2,000$	$1,000$	500
ALR	4.7413	5.4757	8.1634	10.098
MLP	**4.6754**	5.3829	**5.9429**	**6.9916**
SVR	6.4505	7.1429	7.3561	8.5213
Boost	4.9603	**5.2235**	7.7331	7.9125
AVE	5.2064	5.8063	8.0489	8.3809

Table 4. MAE of estimators learned from different training set.

The result displays that, accuracy of estimation increases with decrease of training set size. With 4,000 samples for training, estimation obtains best result with MAE of 5.21 years averagely. When trained from 500 samples, the estimation error is a little higher, with MAE about 8.38 years. But it is also promising considering only 6 images for each age and simple regression methods are adopted. As the design of feature set is guided by the hierarchical model, the features exclude a large amount of redundant features and is of low dimension, which can alleviate over-fitting problem to some extent and obtain good results on small training set.

6. Discussion and future works

This paper proposes an age estimation framework based on a hierarchical face model, guided by which we design a sparse feature set for age estimation. The features take advantages of low dimension, high discriminative power and attains promising performance over previous approaches.

This framework also applies to recognition of other face information, including gender, race, et al. Building a unified system for automatic recognition of facial information will be the focus of our future work.

7. Acknowledgment

This work is done at the Lotus Hill Institute and is partially supported by National Natural Science Foundation of China(NSFC) under contract No.60728203, 100 Talents Program of CAS, ISVISION Technology Co. Ltd and an international science and technology cooperation key-project No.2006CA024,

The data used in this paper is provided by the Lotus Hill Annotation project[17] , which is supported by the NSFC Contact No. 60672162.

References

[1] V. Bruce, et al. "Further experiments on the perception of growth in three dimensions", *Perception and Psychophysic*, 1989. 1, 4
[2] W. B. Horng, et al. "Classification of Age Groups Based on Facial Features", *Tamkang Journal of Science and Engineering*, Vol. 4, No. 3, pp. 183-192, 2001. 1, 4
[3] Y. Kwon and N. D. Vitoria Lobo, "Age Classification from Facial Images", *CVIU*, 1999. 1, 4
[4] D. Kalamani and P. Balasubramanie "Age Classification using Fuzzy Lattice Neural Network", *Intelligent Systems Design and Applications*, 2006. 1, 4
[5] T. Cootes, G. Edwards, and C. Taylor, "Active appearance models", *PAMI*, Vol. 23, No, 6, pp. 681-685, 2001. 1
[6] A. Lanitis, C. Draganova, C. Christodoulou "Comparing Different Classifiers for Automatic Age Estimation", *IEEE Trans. SMC*, Vol. 34, No. 1, Feb. 2004. 1, 5
[7] X. Geng, et al., "Automatic Age Estimation Based on Facial Aging Patterns", *PAMI*, Vol. 29, No. 12, Dec. 2007. 1, 5
[8] Y. Fu, et al. "Human Age Estimation with Regression on Discriminative Aging Manifold", *Multimedia*, 2007. 1
[9] S. Yan, et al. "Learning Auto-Structured Regressor from Uncertain Nonnegative Labels", *ICCV*, 2007. 1, 5
[10] "Face and Gesture Recognition", FG-NET aging database. http://sting.cycollege.ac.cy/ alanitis/fgnetaging/. 1, 4
[11] K. Ricanek, et al. "MORPH: A Longitudinal Image Database of Normal Adult Age-Progression", *FG,* 2006. 1
[12] Z. J. Xu, et al., "A high resolution grammatical model for face representation and sketching", *CVPR,* 2005. 2
[13] J. L. Suo, F. Min, et al. "A Multi-Resolution Dynamic Model for Face Aging Simulation", *CVPR,* 2007. 2
[14] N. Ramanathan and R. Chellappa, "Modeling age progression in young faces", *CVPR*,2006. 4
[15] L. Liang, et al. "Accurate Face Alignment using Shape Constrained Markov Network" *CVPR*, 2006. 4
[16] P. Viola and M. Jones, "Robust Real-time Object Detection", *IJCV*, Vol. 57 , Issue. 2, pp. 137-154, May 2004. 5
[17] Z. Y. Yao, et al., "Introduction to a Large-Scale General Purpose Ground Truth Database: Methodology, Annotation Tool and Benchmarks", *EMMCVPR*,2007. 6