Face Classification Based on Natural Features and Decision Tree

Lingkun Luo*, Shiqiang Hu*, Jiyuan Cai, Fuhui Tang, Zhoujingzi Qiu, Xing Hu,
School of Aeronautics and Astronautics
Shanghai Jiao Tong University
Shanghai, China
lolinkun1988@sjtu.edu.cn, sqhu@sjtu.edu.cn

***Abstract*- Existing face recognition methods suffer from efficiency problems and heavily rely on proper features extraction. In this paper, we propose an efficient face classification method which aims to reduce sensitivity to facial variations and occlusions, meanwhile complete tasks efficiently. In contrast with most energy minimizing based recognition methods, proposed algorithm is cast as a simple classification in our method. First, preprocess images for enhancing data images prior to computational processing and label parts of images as training data. Then we use Active Shape Model (ASM) to extract robust natural features. After that we categorize features and mark them with different labels. Finally, we learn a discriminative C4.5 decision tree for classification. Our method can efficiently classify face images and robust handle facial variations and occlusions. Extensive experiments are conducted on AR database in order to demonstrate the robustness of proposed method. Quantitative and qualitative results compared with several popular algorithms suggest effectiveness and efficiency of proposed method.**

Keywords—natural feature; active shape model; C4.5 decision tree; face calssification; AR database

I. INTRODUCTION

Research activities and demands in recognition of human faces from still image and video image have increased significantly over the past 40 years[1-3]. Face recognition has becoming the most mature and important research topic, due to increase in security concerns and many other applications[15,16] (biometrics, securities, commercial etc). Face recognition is essential for interpreting facial expressions, intentions, behavior, and human emotions, which offers the most natural way of identification. Besides computer graphics and machine learning communities are increasingly involved in face recognition. Fully automatic face recognition algorithm should consistent of two major parts: (1) face detection and normalization; (2) face identification. In this paper we just focus on the latter one.

In recent decades, several popular algorithms for face recognition have been proposed, they are mainly based on eigenfaces[4,5] which argues that any face images could be approximately reconstructed by a standard face (eigenface) and a small collection of weights for other faces. Sparse coding[2,6,7] tries to explain that face belong to the group which contain lest reconstructed error. Fisherface[8,9] shapes the scatter in order to make it more reliable for classification. Laplacianfaces[10] uses locality preserving projections which is the optimal linear approximations to the eigenfunctions on the face manifolds. Deep learning[11,12] recognize face through non linear processes which achieves the best performance. Geometrical feature matching[13,14] computes a set of geometrical features from the picture of face. Most methods mentioned above strongly depend on proper features extraction. Various feature descriptors can be used to represent face images. For example, Haar-like rectangular features, histogram of oriented gradients (HoG) and local binary patterns (LBP).

However, some import weaknesses still exist in recent face recognition methods, they are such as: (1)Some machine learning based methods can achieve good recognition results however they are unable to meet the speed and memory efficiency requirements. (2) Extracted features are mainly based on statistical classification which sensitive to facial variations and facial occlusions. (3) Linear methods are unable achieve good performances on complex data distribution, meanwhile non-linearly methods suffer from the overfitting problem.

II. PROPOSED METHOD

We propose a novel method based on ASM[17,18] and C4.5[19] decision tree (ASM-C4.5) which extracted natural features for faces classification. In order to solve existing limitation of speed and memory issues, we introduce C4.5 algorithm into face classification. Decision tree(DT) presents several advantages[20,21] over other complex learning algorithms, such as low computational cost, robustness to noise, and able to deal with redundant attributes. One of the beauties of DT is it follows human logic and simulates human learning. Furthermore, it takes advantage of BOOST technology to handle large-scale database faster than popular machine learning technologies. These advantages make DT well competitive with those machine learning techniques in literature.

Proposed method classifies faces into different groups. The flowchart of the proposed ASM-C4.5 method is shown in Fig.1. We start by classifying preprocessed face images into training and testing groups as shown in Fig.1a. Then, we extract distances between landmarked points on faces through ASM algorithm and choose discriminative and robust ratios as natural features. In Fig.1d, we categorize natural features into different groups for C4.5 decision tree induction. Finally, we design C4.5 decision tree for classification based on training dataset. Classification results are shown in Fig1.f.

978-1-5090-5188-5/16 $31.00 © 2016 IEEE
DOI 10.1109/ICVRV.2016.10

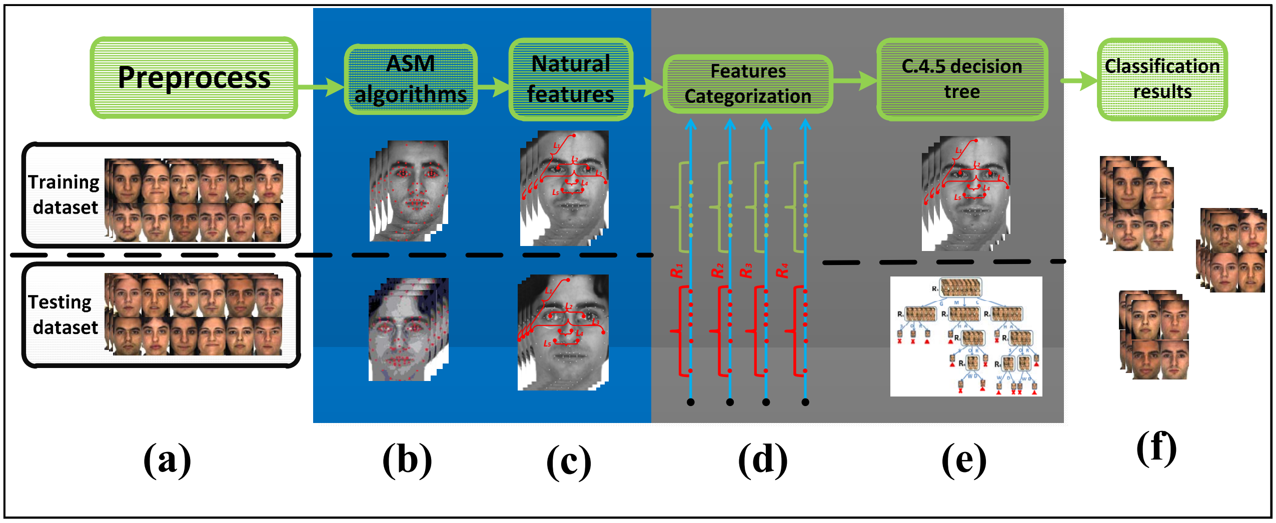

Figure1. The flowchart of the proposed ASM-C4.5 method

The key contributions of our method are as follows:

- Natural features extracted with ASM algorithm are distinctively for classification. In this paper, extracted natural features are less sensitive to facial variations and occlusions.
- ASM-C4.5 is based on the principle of classification which classifies face images efficiently and effectively. Precise face recognition work will benefit from this method.
- Experiments are carried out on the AR[22] database. Results show that ASM-C4.5 displays very competitive results with a number of popular methods. Extensive experiments carried out on different facial variations and occlusions show robustness of extracted natural features.

The remainder of this paper is organized as follows. In Section 3, we introduce the process of natural features extracted with ASM method and the detail of classification process. In Section 4, we compare proposed method with several popular algorithms and discuss robustness and limitation of proposed method. Finally the conclusion and hints to future works are given in Section 5.

III. NATURAL FEATURES EXTRACTION AND C4.5 DECISION TREE FOR FACE CLASSIFICATION

Our proposed ASM-C4.5 method aims at classifying face images into meaningful partitions. As shown in Fig.1, there are three major steps: 1) extract discriminative and robust natural features then label those features (as seen in Fig.1b and Fig.1c); 2) design C4.5 decision tree for classification based on extracted natural features of training dataset(as seen in Fig.1d and Fig.1e); and 3) classification results(as seen in Fig.1f).

A. Natural features extraction based on Active shape model

Here we propose to locate landmarks in images of human faces through ASM[17] algorithm. Then, we extract natural features for C4.5 decision tree induction.

1) Definition of natural features

The human face conveys to other human beings, and potentially to computer systems, such information as intentions, gender, age and ethnicity. However human beings and computer systems share different ways of understanding human faces. The computer approaches[4,9,23] to understand face can be roughly divided into holistic and feature based methods. Holistic techniques perform an automatic extraction of data, on the basis of some general rule (LDA, PCA, Wavelets, etc.), meanwhile feature based approach always combine with statistical classification.

Human beings try to understand face images using facial features, and geometric features. We are insensitive to some features (e.g., nose width, intraocular distance, size of eyes, shape of face). Some researches[24-26] show scientists and artists have debate the definition of beauty, which contain the golden ratio, geometric features, etc. Those features offer a non-intrusive, and perhaps the most natural way of identification. Therefore, we define those features as natural features.

The natural features extracted in this paper are the ratio between the feature point distances. The natural feature defined in this paper tries to simulate human understanding and capture attention of faces. Meanwhile those features are robust to facial variations and occlusions

2) ASM method for landmarking

ASM models have been around in computer vision research for more than 20 years. They are statistical models for locating landmarks in images of human faces. Proposed ASM algorithm can capture natural features (e.g., nose width, intraocular distance).

Mask should be defined first in order to capture feature points on face image precisely. This paper propose ASM algorithm to extract face feature points.

The main steps of the ASM algorithm include:

a) Mathematical expression of feature points.

Mask X_i can be expressed by a series of longitudes and latitudesas:

$$X_i = (x_{i,1}, y_{i,1}, x_{i,2}, y_{i,2}, \cdots x_{i,j}, y_{i,j})^T, \qquad (1)$$

where (x_{ij}, y_{ij}) is the coordinates of *j*th feature points in *i*th image, i=1,2, ⋯N, j=1,2, ⋯,77. N is the number of images in database, n is the number of feature points existed in each face image(n=77).

b) Image alignment

Suggest a new shape by adjusting the current positions of the landmarks. After rotate and resize image we get the minmal E^2 as:

$$E^2 = (X_1 - M(s,\theta)[X_2] - t)^T w(X_1 - M(s,\theta)[X_2] - t), \qquad (2)$$

where $M(s,\theta)[X_2]$ represent θ degrees of X_2 rotation and resize X_2 with ratio s as:

$$\begin{aligned} M(s,\theta)[X_2] &= \begin{pmatrix} s\cos\theta & -s\sin\theta \\ s\sin\theta & s\cos\theta \end{pmatrix} \begin{pmatrix} x_{jk} \\ y_{jk} \end{pmatrix} \\ &= \begin{pmatrix} (s\cos\theta)x_{jk} - (s\sin\theta)y_{jk} \\ (s\sin\theta)x_{jk} + (s\cos\theta)y_{jk} \end{pmatrix}, \end{aligned} \qquad (3)$$

Finally, the landmark is moved to the center of the patch which best matches the landmark's model descriptor

c) Modeling

With extracted linera combinations from aligned shape through PCA model, then we built subspace descrbes shape features. Any face can match with ASM model with built subspace.

Feature points will be extracted from ASM model after the above steps. Fig.2 shows the extracted 77 feature points on face image defined in this paper. The distributions of those points are list as follows: 3 feature points exist on forehead, 6 feature points on left eyebrow, 6 feature points on right eyebrow, 10 feature points on left eye, 10 feature points on right eye, 11 feature points on nose, 18 feature points on mouth and 13 feature points on contour of face.

3) Extract discriminative and robust natural features

Euclidean distance of each two feature points as:

$$D_{(A,B)} = \sqrt{(A_X - B_X)^2 + (A_Y - B_Y)^2}, \qquad (4)$$

where A and B are two feature points as shown in Fig.2. A_X, A_Y is horizontal and vertical coordinates of point A,

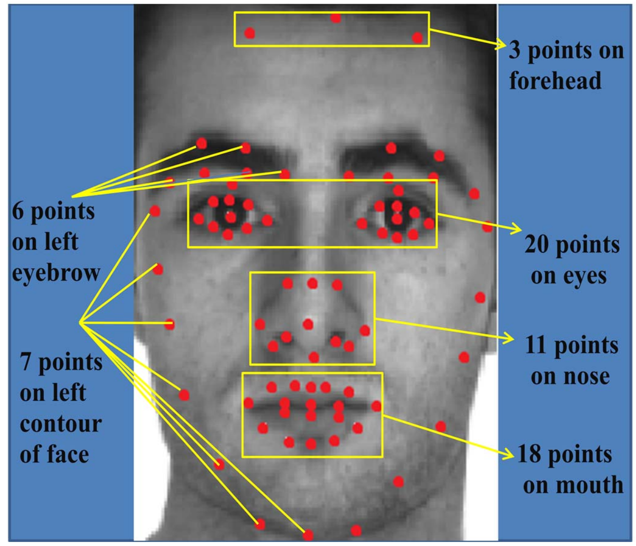

Fig. 2 Extracted 77 feature points

meanwhile B_X, B_Y is horizontal and vertical coordinates of point B respectively. We get 2926 distance D_{2926} from the 77 points on each face.

Fig.3(a,b) show $D_{(A,B)}$ is robust to facial variations occlusions, however, $D_{(C,D)}$ is sensitive to facial variations and facial occlusions. In order to select robust natural features and understand facial images we choose stable and robust distances through calculate the variance of each distance of facial images from one people in AR database:

$$e_{ki} = \frac{[(\frac{d_{ki1} - m_{ki}}{m_{ki}})^2 + (\frac{d_{ki2} - m_{ki}}{m_{ki}})^2 \cdots + (\frac{d_{kij} - m_{ki}}{m_{ki}})^2]}{j}, \qquad (5)$$

where $i = 1,2,\ldots 2926$, $j = 1,2,\ldots,26$ and $k = 1,2,\ldots 100$. d_{kij} is the i th distance of d_{2926} in j th faces of k th people, m_{ki} is the i th average distance of k th people, and e_{ki} is the is the i th distance variance of k th people. Here we propose database contains over 2000 color face images of 100 people, each people provide 26 facial images. We also calculate the variance of distance of the database proposed:

$$E_i = \frac{[(\frac{D_{i1} - M_i}{M_i})^2 + (\frac{D_{i2} - M_i}{M_i})^2 \cdots + (\frac{D_{ia} - M_i}{M_i})^2]}{a}, \qquad (6)$$

where a is the number of images. D_{ia} the i th distance of d_{2926} in a th image. M_i is the i th average distance of proposed database. The less value the E_i is the more stable the distance is. We select testing images for designing C4.5

tree through comparing e_{ki}. Here we define natural features proposed in our methods as:

$$R_a = \frac{Stable\ d_i}{Stable\ d_j}, \tag{7}$$

where $(d_i, d_j) \in D_{2926}$,and R_a is the natural feature which robust to facial variations and occlusions. We choose stable R_a for designing C4.5 tree through same theory as equation(6).

To better understand natural features. We choose five relatively stable distances for creating natural features as Fig.3c shows. D_1 is the distance between forehead and right cheekbone, D_2 is the distance between two eyes, D_3 is the cheekbone width, D_4 is the width of the bridge of the nose, D_5 is the nose width. Finally we choose four stable natural features:

$$R_1 = D_2 / D_4 \tag{8}$$

$$R_2 = D_3 / D_2 \tag{9}$$

$$R_3 = D_5 / D_4 \tag{10}$$

$$R_4 = D_3 / D_1 \tag{11}$$

Selected R_1, R_1, R_1 and R_1 are parts of stable natural features.

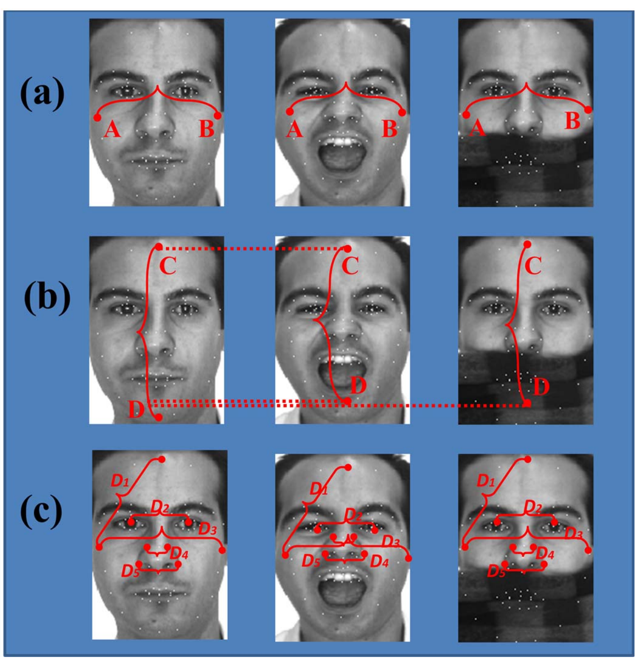

Fig.3 Selected distances (a) Stable distance, (b) Unstable distance (c) Extract stable distances

B. Decision tree classify method

Decision tree algorithm is one of the most popular classification algorithms. It does not require users to understand detail background knowledge in the learning process, as long as the training samples can be expressed by using attribute values. DT classification deals with three tasks, i.e. the choice of feature subsets to be used at each internal node, the choice of the decision rule and the choice of decision tree structure.

C4.5/C5.0 based on ID3 is the most popular decision tree algorithms, which not only classify more than one categories and allow continuous value data to study, but also convert decision tree to equivalent production rules. Here we propose C4.5 DT for classification

1) Prerequisites for decision tree building

In this paper, we use natural features extract from AR database to design our C4.5 decision tree. We choose 294 images from database for decision tree induction. Value of each feature is not suitable for mining meaningful decision making rules, so we categorize each feature into different attributes.

We propose K-means algorithm to classify four groups one dimension vector of R_1, R_2, R_3 and R_4 respectively. After analysis those four vector, we categorize R_1, R_2, R_3 and R_4 data into 3,3,2,2 groups respectively. Here we show the main processes:

Input: Groups number k, and n natural feature value.

Output: k groups natural feature

Step(1): Assign initial value for k classes center.

Step(2):For $j=1$ to n assign each x_j to the closest center.

Step(3):For i=1 to k compute the average of x_i and variance of E.

$$\bar{x}_i = \frac{1}{|C_i|} \sum_{x \in C_i} x \tag{12}$$

$$E = \sum_{i=1}^{k} \sum_{x \in C_i} |x - \bar{x}_i|^2 \tag{13}$$

Step(4):Set the $\bar{x}_i$ as k new center, repeat step 2.

Step(5):Stop until E convergent.

The stopping criterions in our algorithm is E become stable. The categorize results for R_1, R_2, R_3, and R_4 are shown in TABLE I. In TABLE I the different attributes of selected natural features are defined as 'O', 'S', 'R', 'G', 'C', 'M', 'H', 'A', 'W' and 'D'. Each attribute represents different range of feature value.

TABLE I. ATTRIBUTES EXTRACTION

Natural feature	Attributes		
R_1	O(Value>3.23529)	S(rest)	R(Value <3.05882)
R_2	G(Value >2.12723)	C(rest)	M(Value <2.03644)
R_3	H(Value >2.0086)		A(Value <2.0086)
R_4	W(Value >1.13209)		D(Value <1.13209)

2) Process of C4.5 decision tree

C4.5 builds decision tree from a set of training data similar as ID3. C4.5 chooses the attributes of the data which most effectively splits its set of samples into subsets enriched in one class or other. Proposed C4.5 selects attribute with largest gain ratio as split attribute which design more efficient decision tree.

$$ENT(D_J) = -\sum_{i=0}^{c-1} p_i \log_2 p_i \tag{14}$$

$$SPLIT\_ENT(D_J) = -\sum_{i=0}^{c-1} p_i ENT(D_J) \tag{15}$$

where c represents the number of attributes and ENT is entropy. Each attribute will divide dataset into two or more sub-datasets. The gain ratio of this attribute is defined in (10).

$$GAIN\_RATIO(A) = \frac{ENT(D) - \sum_{j=1}^{k} \frac{|D_j|}{|D|} ENT(D_j)}{SPLIT\_ENT(A)} \tag{16}$$

A is the attribute candidate, k is branch number of attribute, D is the original dataset, D_j is the sub-dataset divided by A.

Finally we choose the attribute with largest value of GAIN_RATIO as split attribute.

IV. EXPERIMENTS AND DISCUSSION

In this section, we introduce the details about the test database, measurements and experiment environment. We compare proposed method under different facial appearances. We also report the evaluation results of proposed ASM-C4.5 by comparing with other methods on AR database.

A. Database and evaluation metrics

The database used in this paper was obtained from AR database. Our experiment environments are on a PC with 3.2-G CPU, 4.0G RAM and Matlab 2012.

1) ProposedDatabase

The AR face database contains over 4000 colorful face images of 126 people, including frontal views of faces with different facial occlusions, lighting conditions, and facial expressions. We manually cropped the face portion of image and then normalized it to 60*80 pixels. After analysis AR database we categorize it into 7 groups as shown in TABLE II.

TABLE II. CIRCUMSTANCE COMBINATION

Face Group No.	Varity circumstance			
	Facial variations	*Illumination variations*	*Sunglasses*	*Scarf*
No.1(1,3,14)	N	N	N	N
No.2(2,4,15,16,17)	Y	N	N	N
No.3(5,6,7,18,19,20)	N	Y	N	N
No.4(8,21)	N	N	Y	N
No.5(11,24)	N	N	N	Y
No.6(9,10,22,23)	N	Y	Y	N
No.7(12,13,25,26)	N	Y	N	Y

In AR database each person provide 26 facial images so No.1(1,3,14) means 1st, 3rd, 14th facial images categorized into face group No.1. Proposed AR database contain facial images with different situations. Facial variations, illumination variations, and occlusions will affect ASM algorithm seriously, however, proposed natural feature is robust to illumination variations and occlusions. After analysis ASM results we choose group No.1 and No.4 for designing C4.5 decision trees.

2) Evaluation metrics

We propose recognition rate (e) for quantitative comparison. The recognition rate is defined as ratio of the number of correctly classified face images to the number of the total test face images.

$$e = \frac{no.correctly\ classified\ faces}{no.total\ test\ faces} \tag{17}$$

B. C4.5 Decision tree

We choose 294 images of AR database from group No.1 and No.4 as TABLE II mentioned. Those 294 facial images labeled as learning database. Extract attributes of learning database are list in TABLE III.

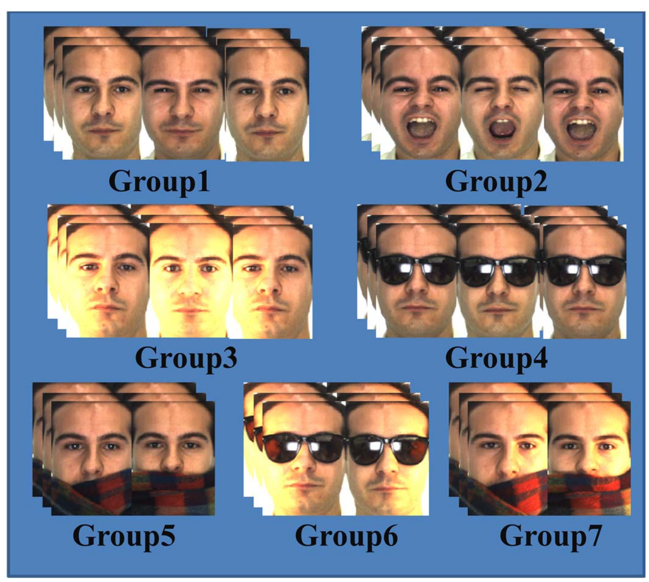

Fig. 4 AR database divided into seven groups.

TABLE III. EXTRACT ATTRIBUTES

Attribute Group No.	Attributes			
	R_1	R_2	R_3	R_4
No.1	R	C	A	W
No.2	O	M	H	W
No.3	R	G	A	W
No.4	R	M	H	D
No.5	R	C	A	D
No.6	S	M	A	D
No.7	O	M	A	D
No.8	O	M	H	D
No.9	S	C	A	W
No.10	S	M	H	D
No.11	O	C	A	D
No.12	S	C	H	W
No.13	R	M	A	D
No.14	O	C	H	W
No.15	S	C	A	D
No.16	O	M	A	W
No.17	O	G	A	W
No.18	O	C	A	W
No.19	R	C	H	D
No.20	S	C	H	D
No.21	R	M	A	W

We classify those 21 attribute groups list in TABLE III. In order test robustness of proposed method we design two C4.5 decision trees based on TABLE III for experiments. C4.5 decision tree No.1 is design based on attributes No.1 to No.10. C4.5 decision tree No.2 is design based on attributes No.11 to No.21. The C4.5 decision tree No.1 is as Fig.5 shows. The attributes is not redundant so there is no pruning process.

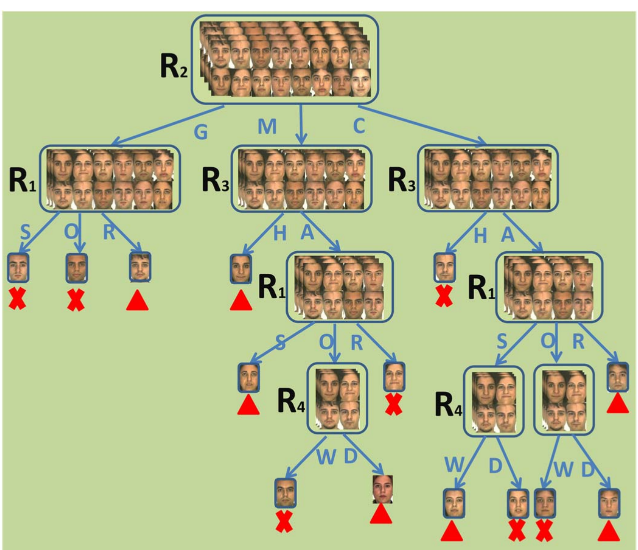

Fig. 5 No.1 C4.5 Decision tree

On Fig.5 the leaf node which marked red 'X' means faces not belong to the decision tree and red triangle means the face contain natural features above. We also build No.2 C.5 decision tree follow the same processes.

C. Quantitative comparision

In this section we quantitatively compare results of two C4.5 decision trees with facial images in different conditions, meanwhile compare ASM-C4.5 with other popular methods. All comparisons are conducted on AR database with same evaluation metrics.

1) Comparision between different conditions

We compare results between two C4.5 decision trees. TABLE IV lists the recognition rates of two decision trees with different facial groups (facial variations, occlusions, illumination variations). Quantitatively comparison results between ASM-C4.5 with PCA, LDA and KNN are shown in TABLE V.

From TABLE IV, it can be seen the two C4.5 decision trees get almost similar results. They get lowest recognition rate in face group No.3 and No.7 which means strong illumination variations and illumination variation combine with facial occlusion affect results seriously.

TABLE IV. COMPARISON RESTLTS BETWEEN TWO C4.5 DECISION TREES

Results Group No.	Test results	
	Tree No.1	*Tree No.2*
No.1(1,3,14)	95.4%	95.5%
No.2(2,4,15,16,17)	90.4%	88.5%
No.3(5,6,7,18,19,20)	75%	72.4%
No.4(8,21)	94.2%	93.8%
No.5(11,24)	90.8%	90.5%
No.6(9,10,22,23)	72.8%	78.2%
No.7(12,13,25,26)	71.3%	61.6%
Total results	89.6%	86.4%

2) Comparision between different methods

We also compare proposed method with some popular face recognition methods as TABLE V shows.

TABLE V. COMPARISON RESTLTS BETWEEN DIFFERENT ALGORITHMS

Test results				
Tree No.1	*Tree No.2*	*PCA*	*LDA*	*K-nearest neighbors*
89.6%	86.4%	82.5%	70.5%	58.4%

TABLE V shows our proposed method achieves highest recognition rate, and outperform some popular methods.

D. Discussion

In this section we discuss robustness and weakness of proposed method.

- The two C4.5 decision tree achieve nearly similar results shows proposed method is robust. However they both achieve not so good results in face groups No.3 and No.7 represent ASM-C4.5 is sensitive to strong illumination variations. This is mainly because

ASM algorithm failed on strong illumination variations which lead to bad natural features.

- TABLE V shows PCA and LDA achieve poor results. That's because PCA is unsupervised method and there is facial occlusions and variations in AR database meanwhile the amount of learning samples is not sufficient. These factors lead the linear method does not work so well. LDA is unable to handle the classification mainly because face database contain homogeneous data distribution which can't attribute as a two class problem
- Our method is more efficient than some popular machine learning algorithms. We strong believe with more training data we can learn a more robust and distinctive decision tree for classification

V. CONCLUSIONS/FUTURE WORK

In this paper, a face classification approach combining ASM algorithm with C4.5 decision tree classification is proposed. The method extracts natural features, which is robust to facial variations and occlusions. Additionally the classification strategy used in this paper can efficiently classify face images. Experiments on AR database demonstrate that ASM-C4.5 can obtain satisfactory results comparing with some popular algorithms.

Future work will concentrate on extracting more robust natural features for C4.5 decision tree in order to handle strong light illuminations and design more robust decision tree algorithms.

ACKNOWLEDGEMENTS

This paper is jointly supported by the National Natural Science Foundation of China "62174161", China Aviation Science Foundation "20142057006". We thank the reviewers and editors for their comments and suggestions.

REFERENCES

[1] Zhao W, Chellappa R, Phillips P J, et al. Face recognition: A literature survey[J]. ACM computing surveys (CSUR), 2003, 35(4): 399-458.

[2] Wright J, Yang A Y, Ganesh A, et al. Robust face recognition via sparse representation[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2009, 31(2): 210-227.

[3] Yang J, Zhang D, Frangi A F, et al. Two-dimensional PCA: a new approach to appearance-based face representation and recognition[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2004, 26(1): 131-137..

[4] Turk M, Pentland A. Eigenfaces for recognition[J]. Journal of cognitive neuroscience, 1991, 3(1): 71-86.

[5] Mohammed A A, Minhas R, Wu Q M J, et al. Human face recognition based on multidimensional PCA and extreme learning machine[J]. Pattern Recognition, 2011, 44(10): 2588-2597.

[6] Zhang L, Yang M, Feng X. Sparse representation or collaborative representation: Which helps face recognition?[C]//Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, 2011: 471-478.

[7] Wagner A, Wright J, Ganesh A, et al. Toward a practical face recognition system: Robust alignment and illumination by sparse representation[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2012, 34(2): 372-386.

[8] Martínez A M, Kak A C. Pca versus lda[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2001, 23(2): 228-233.

[9] Jing X Y, Wong H S, Zhang D. Face recognition based on 2D Fisherface approach[J]. Pattern Recognition, 2006, 39(4): 707-710.

[10] He X, Yan S, Hu Y, et al. Face recognition using Laplacianfaces[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2005, 27(3): 328-340.

[11] Sun Y, Wang X, Tang X. Deep learning face representation from predicting 10,000 classes[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014: 1891-1898.

[12] Fan H, Cao Z, Jiang Y, et al. Learning deep face representation[J]. arXiv preprint arXiv:1403.2802, 2014.

[13] Kotsia I, Pitas I. Facial expression recognition in image sequences using geometric deformation features and support vector machines[J]. Image Processing, IEEE Transactions on, 2007, 16(1): 172-187.

[14] Tan X, Triggs B. Enhanced local texture feature sets for face recognition under difficult lighting conditions[J]. Image Processing, IEEE Transactions on, 2010, 19(6): 1635-1650.

[15] Kong S G, Heo J, Abidi B R, et al. Recent advances in visual and infrared face recognition—a review[J]. Computer Vision and Image Understanding, 2005, 97(1): 103-135.

[16] Jafri R, Arabnia H R. A Survey of Face Recognition Techniques[J]. JIPS, 2009, 5(2): 41-68.

[17] Milborrow S, Nicolls F. Active Shape Models with SIFT Descriptors and MARS[C]//VISAPP (2). 2014: 380-387.

[18] Milborrow S, Nicolls F. Locating facial features with an extended active shape model[M]//Computer Vision–ECCV 2008. Springer Berlin Heidelberg, 2008: 504-513.

[19] Chawla N V. C4. 5 and imbalanced data sets: investigating the effect of sampling method, probabilistic estimate, and decision tree structure[C]//Proceedings of the ICML. 2003, 3.

[20] Wang X Z, Dong L C, Yan J H. Maximum ambiguity-based sample selection in fuzzy decision tree induction[J]. Knowledge and Data Engineering, IEEE Transactions on, 2012, 24(8): 1491-1505.

[21] Barros R C, Basgalupp M P, De Carvalho A C, et al. A survey of evolutionary algorithms for decision-tree induction[J]. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 2012, 42(3): 291-312.

[22] Martinez A M. The AR face database[J]. CVC Technical Report, 1998, 24.

[23] Shen L, Bai L. A review on Gabor wavelets for face recognition[J]. Pattern analysis and applications, 2006, 9(2-3): 273-292.

[24] Laurentini A, Bottino A. Computer analysis of face beauty: A survey[J]. Computer Vision and Image Understanding, 2014, 125: 184-199.

[25] Aarabi P, Hughes D, Mohajer K, et al. The automatic measurement of facial beauty[C]//Systems, Man, and Cybernetics, 2001 IEEE International Conference on. IEEE, 2001, 4: 2644-2647.

[26] Pallett P M, Link S, Lee K. New "golden" ratios for facial beauty[J]. Vision research, 2010, 50(2): 149-154.