Design and Implementation of Effective Pattern Classification Model for Face Recognition

Do-Hyeon Kim [1], Eui-Young Cha [2], Kwang-Baek Kim [3]
[1,2]Department of Computer Engineering, Pusan National University, Busan, Korea
[3] Department of Computer Engineering, Silla University, Busan, Korea
e-mail : unlimmit@hanmail.net, eycha@pnu.edu, gbkim@silla.ac.kr

Abstract— **This paper presents an effective pattern classification model by designing an artificial neural network based pattern classifiers for face recognition. First, a RGB image inputted from a frame grabber is converted into a HSV image which is similar to the human beings' vision system. Then, the coarse facial region is extracted using the hue(H) and saturation(S) components except intensity(V) component which is sensitive to the environmental illumination. Next, the fine facial region extraction process is performed by matching with the edge and gray based templates. To make an light-invariant and qualified facial image, histogram equalization and intensity compensation processing using illumination plane are performed. The finally extracted and enhanced facial images are used for training the pattern classification models. The proposed H-ART2 pattern classification model which has the hierarchical ART2 layers make it possible to search clustered reference patterns effectively. Experimental results show that the proposed face recognition system is as good as the SVM model which is famous for face recognition field in recognition rate and even better in classification speed.**

.

I. Introduction

Recently, face recognition becomes the important research area for natural human-computer interaction. The Process of face recognition is generally composed of facial region extraction, feature extraction from acquired facial images, and recognition from extracted features. Facial region extraction is the technique which automatically extracts the position of face from a image. In early days, facial region extraction is the stage of preprocessing for face recognition. But it becomes the independent research area due to its importance. There are various techniques of facial region extraction: skin color based method using color information[1], motion-information method, template matching, neural networks method[13], and snake method. The recognition methods of extracted facial images are largely classified into geometrical feature matching, statistical method[16], PCA(Principle Component Analysis) method[12], LFA(Local Feature Analysis) method, LDA(Linear Discriminant Analysis) method[3], neural networks method[11,14,15] and SVM(Support Vector Machine) method[1,9].

This paper presents an effective pattern classification model by designing an artificial neural network based pattern classifiers for face recognition. First, a RGB image inputted from a frame grabber is converted into a HSV image. Then, the coarse facial region is extracted using the hue(H) and saturation(S) components except intensity(V) component. Next, the fine facial region extraction process is performed by matching with the edge and gray based templates. To make a light-invariant and qualified facial image, histogram equalization and intensity compensation processing are performed. The finally extracted and enhanced facial images are used for training the pattern classification models. The proposed H-ART2 pattern classification model which has the hierarchical ART2 layers make it possible to search clustered reference patterns effectively

II. Preprocessing and Facial Region Extraction Technique

A. Coarse Facial Region Extraction by HSV Color Model

The RGB image inputted from a frame grabber is converted into a HSV image which is similar to the human beings' vision system. Then, the region of which color is similar to facial color is extracted using the hue and saturation components except intensity component which is sensitive to the environmental illumination. Namely, the distribution of hue and saturation components corresponding to color of human facial skin by experiential value is obtained. And then, hue and saturation mask is made from the corresponding region of the image with a certain threshold. After the common region is extracted from two masks, the final facial skin mask is generated by interpolation to obtain the complete facial skin mask. The coarse facial region including a face is extracted as the circumscribed region of that facial skin region mask

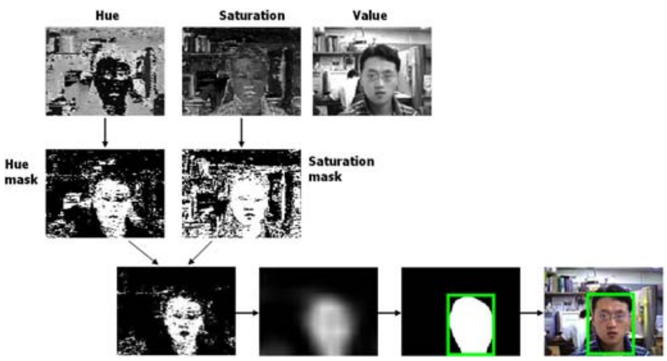

Fig. 1. Coarse facial region extraction by color information

B. Fine Facial Region Extraction by Dual Templates Matching based on Correlation

a. Creating Dual Templates

50 * 50 gray-based and edge-based template are used to minimize the interference of illumination. After facial regions is manually extracted from various human face images, both templates are created from the average image of these facial regions. To minimize the interference of illumination and to get balanced templates, right-left mirror images are added. And the gray template is finally equalized.

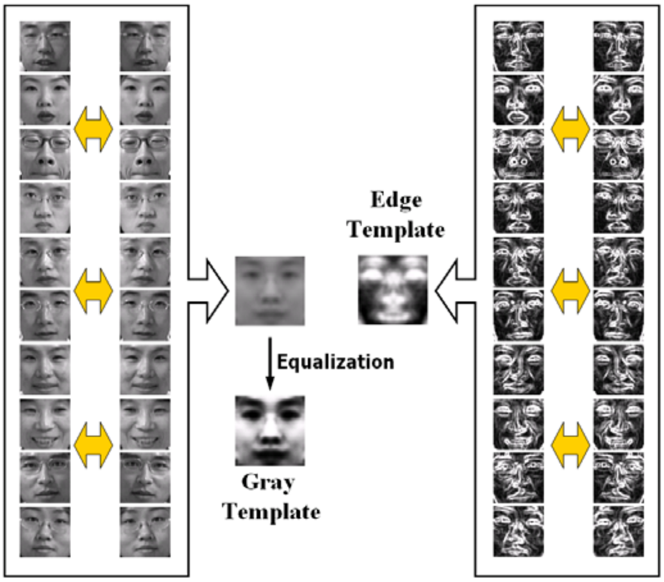

Fig. 2. Creating dual templates.

b. Templates Matching by Correlation

To improve the matching speed, we execute the matching with only facial region which is acquired by color information. To eliminate the effect of template size, we use various scaled versions of original image. The matching original image with template is performed by correlation. Given template T and image window R in original, correlation coefficient $\gamma(T,R)$ is computed by equation (1).

$$\gamma(T,R) = \frac{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}(T[i][j]-\mu_T)(R[i][j]-\mu_R)}{M \bullet N \bullet \sigma_T \bullet \sigma_R} \qquad (1)$$

In equation (1), M, N are the size of image window, μ_T, μ_R are average of T,R, and σ_T, σ_R are the standard deviation of T,R.

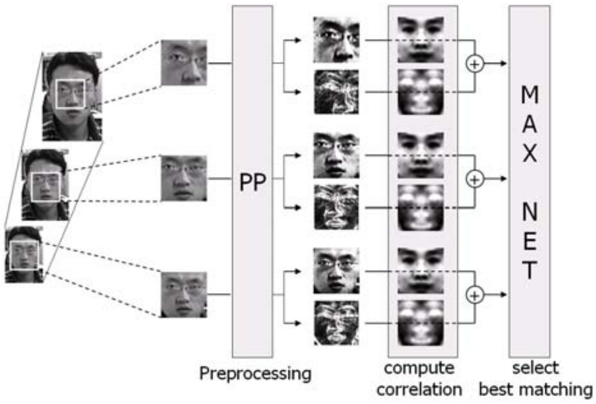

Fig. 3. Templates matching by correlation

After template matching gray template and edge template with image window region in original image are executed, the average of both correlation coefficients is acquired. The region with the largest average correlation coefficient is set up as the optimal facial region in the current scale. This process is repeated at various scales 0.6 – 1.5. Then, the region with the highest correlation coefficient is determined as the final facial region.

C. Image Enhancement by Illumination Plane and Equalization

Generally histogram equalization is used to improve an unequal intensity distribution. However, equalization makes a dark or bright image to have equal intensity distribution in the whole. When the effect of intensity is added to image, it may result in a maximizing illumination effect and have a bad influence on recognition performance. We have designed illumination plane to compensate the effect of illumination.

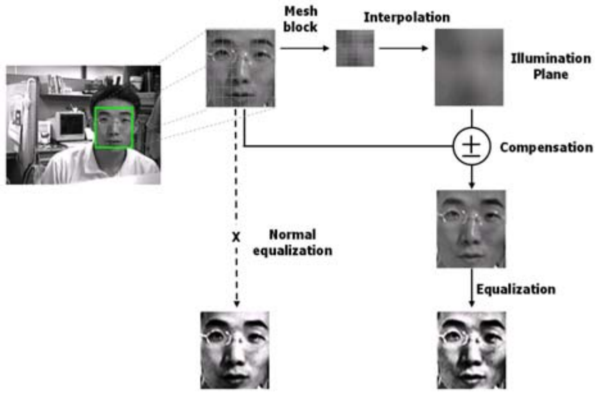

Fig. 4. Image enhancement by illumination plane and equalization

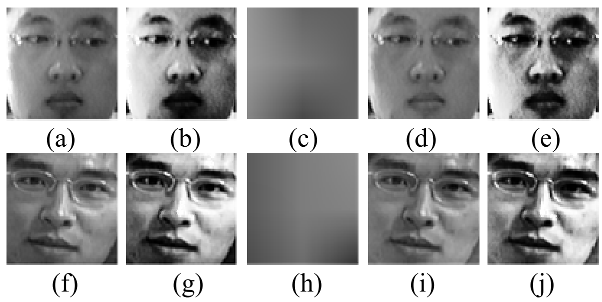

Fig. 5. Comparison of image enhancement by illumination plane and equalization

(a),(f) facial region extraction, (b),(g) equalization, (c),(h) illumination plane,

(d),(i) illumination compensation (e),(j) equalization after compensation

Illumination plane can divide the extracted facial region into blocks, and adjust compensative amounts of illumination according to the size of blocks. In other words, after the representative intensity values of blocks are extracted, illumination plane image can be generated by interpolating to the size of original image

III. PATTERN CLASSIFICATION MODEL

A. Structure of Hierarchical

The structure of ART2(Adaptive Resonance Theroy) is determined by vigilance parameter ρ. It defines the similarity of the stored cluster patterns and a certain input pattern. Or, As ρ is large, the diameter of a cluster is large and the cluster can accept large patterns. As ρ is small, the diameter of a cluster is small, many clusters can be generated, and various patterns can be stored. But because the number of patterns is too many, it may result in lower recognition speed. In this paper, considering such a vigilance parameter, we proposed the Hierarchical ART2 which is consists of two layer clusters.

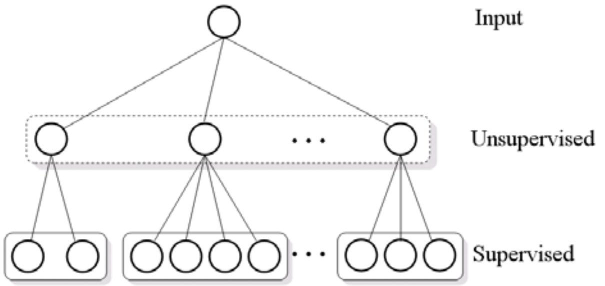

Fig. 6. Structure of H-ART2

H-ART2 is consist of the first clusters connected to the input layer and second clusters respectively connected to the first cluster. The first clusters as a centroid reference vector patterns are the representative patterns for representing input pattern. In this paper, the input patterns classified by each first clusters are more specifically re-classified, and are formed as second clusters.

B. Learning Algorithm of H-ART2

The first clusters H-ART2 are generated by unsupervised learning of the conventional ART2 method. The input patterns of first clusters are more specifically re-classified into the second clusters by the supervised learning algorithm which is a modified ART2 learning algorithm.

The supervised learning algorithm of ART2 is similar to unsupervised learning algorithm. But, although an input pattern may pass the vigilance test, a new cluster is generated in the case that the class of the input pattern is different from the class of winner cluster.

Step 1. x_k is k-th input pattern, w_i is the center value of i-th cluster in neural network, LX_k, LC_i is the label which is representing the respective class of input pattern x_k and the center value of cluster w_i.

Step 2. Select the cluster j* which has the minimum distance to a new input pattern x_k as the winner cluster.

$$\|x_k - \omega_{j*}\| = \min\|x_k - \omega_i\|, (1 \le i \le c) \quad (2)$$

Step 3. Perform the vigilance test of an input pattern. If input pattern exists in the diameter of winner cluster (vigilance parameter), and input pattern and winner cluster exists in the same class, the input pattern is included into the winner cluster and the center value of the cluster is updated. Otherwise, the current input pattern is assigned to a new cluster.

$$\omega_{j*}^{new} = \frac{x_k + \omega_{j*}^{old} \bullet \|Cluster_{j*}^{old}\|}{\|Cluster_{j*}^{old}\| + 1} \quad (3)$$

Input pattern set X={ x_1 , x_2 , …, x_N }
Cluster set C={ w_1 , w_2 , …, w_c }
N: number of input pattern,
c: number of cluster,
T: number of total cycles
for t = 1, 2, 3, … , T
{
for k = 1, 2, 3, … , N
{
if c ==0
c = c+1, w_c= x_k , LC_c=LX_k
else
a.Find winner cluster
$\| x_k - w_{j*} \| = \min \| x_k - w_c \|$, (1<=i<=c)
b.Perform vigilance test
if $\| x_k - w_{j*} \| < \rho$ && LC_{j*}==LX_k
$$\omega_{j*}^{new} = \frac{x_k + \omega_{j*}^{old} \bullet \|Cluster_{j*}^{old}\|}{\|Cluster_{j*}^{old}\| + 1}$$
else
c = c+1, w_c= x_k , LC_c=LX_k
}
if C_t==C_{t-1} then stop
}

Fig. 7. Supervised ART2 learning algorithm

Step 4. Repeat from Step 2 to Step 3 until every input patterns are exhausted.

Step 5. Terminate learning if a designated numbers of learning cycles is repeated or there is no change in the center values of clusters.

C. Searching Strategy

The numbers of clusters in ART2 is determined by vigilance parameter ρ. If the number of clusters is too large, the speed of learning and recognition becomes slow because there are too many clusters to be compared. On the contrary, if too small, the recognition rate becomes low because there is not sufficient space enough to represent various types of input patterns. Considering this trade-off, we have designed a hierarchical neural network structure for recognition performance and have introduced an appropriate method for selecting reference cluster.

Fig. 8. (a) shows the method of selecting only first reference cluster and Fig. 8. (b) shows the method of selecting every second clusters as candidate clusters. Both methods, without searching strategy, select the generated clusters as candidate clusters, and then search for the reference cluster pattern, which is most similar to input pattern, out of them. Other cases of Fig. 8. have their own searching strategies.

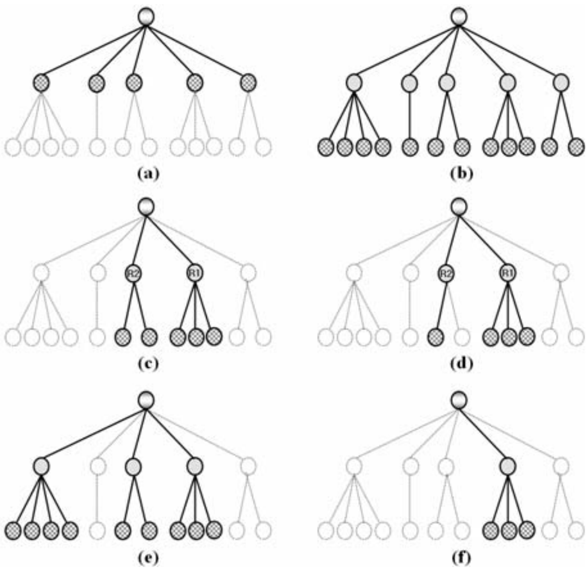

Fig. 8. Searching strategies for candidate clusters.

a. Selection by ranking
In the way of Fig. 8. (c), you obtain the distances between the input pattern and every first clusters, sort the distances, select the size k of most similar first clusters, and finally select each clusters, which connected the previous similar clusters, as candidate clusters. For example, if cluster selection ratio is c and there are the size c of first clusters, the size c x p of first clusters are selected and each clusters, which is connected to the selected, are selected as candidate clusters. An input pattern is classified into the class of the candidate cluster which has the shortest distance between the input pattern and the selected candidate clusters.

b. Selection by 2-WC(Winner Classifier)
In the manner of Fig. 8. (d), you obtain the distances between the input pattern and every first clusters, sort the distances, select only the most similar two first clusters, and finally select every second clusters, which are connected to the most similar one first cluster, and the most similar one cluster, which is connected to the next similar one first cluster, as candidate clusters. Be contrary to selection by raking, an input pattern is classified into the class of candidate cluster which has the most frequent class out of the selected candidate clusters.

c. Selection by Relative Distance Ratio
In the way of Fig. 8. (e) (f), you obtain the distances between the input pattern and every first clusters, calculate the maximum and minimum distances, select only the first clusters, which has the ratio within the predefined distance ratio, and finally select each clusters, which connected the previous selected clusters, as candidate clusters. For instance, if cluster the selection ratio is p, the minimum distance is d_{min} and the maximum distance is d_{max}, only the first cluster, which has the distance below $d_{threshold}$, are selected as candidate clusters. Finally, an input pattern is classified into the class of the candidate cluster which has the shortest distance between the input pattern and the selected candidate clusters.

$$\| w_i - x_k \| \leq d_{threshold} \tag{4}$$

$$d_{threshold} = d_{\min} + (d_{\max} - d_{\min}) \times p \tag{5}$$

Fig. 8. (e) shows that three first reference cluster is similar to the input pattern. Fig. 8 (f) shows that only one reference cluster is similar to the input pattern and that there is the strong points of adaptively selecting reference cluster according to the similarity to the input pattern, not the fixed size of reference clusters.
H-ART2 neural network can improve recognition speed without lowering recognition rate because it selects a few first clusters which is similar to the input pattern according to relative distance ratio, it can compare only the second clusters which are connected to the previous first clusters.

D. Pattern Classification of H-ART2
The first clusters selected by searching strategy are similar to input pattern and, in the same manner, the second cluster connected to the first clusters has similar pattern with input pattern. On the contrary, the unselected first clusters and the second clusters connected to these first clusters don't have similar pattern with input pattern. Therefore fast recognition can be executed because a plenty of clusters has no possibility are pruned. Selecting candidate clusters by cluster selection ratio, not simple similarity with input pattern, can decrease the risk of misclassification.

IV. EXPERIMENTAL RESULT

In this paper, facial data is acquired from the 320 x 240 x 24 (bit color) motion data of CCD camera. The number of people is 10, sex ratio is 6:4, and the ratio of people wearing glasses is 7:3. 1479 facial images of total 2215 is used to the training and the rest 736 is put to the test.

Fig. 9. Facial region data

A. Performance Evaluation of pattern classification model Preprocessings
Facial data are normalized with 25 x 25 and are enhanced with illumination plane and equalization. For extract feature extraction, 3 kinds of image are prepared: original gray images, edge images for emphasizing the outline of face, and combined images from two previous images.

Table 1. Preprocessing of facial input pattern

preprocessing	normalization	enhancement	feature extraction
PF1	25 x 25 normalization	Illumination plane equalization	gray image
PF2			edge image
PF3			gray + edge image

a. Creation of Reference Patterns

Creating reference patterns of H-ART2, according to various vigilance parameters, the change of the number of first clusters and second clusters can be compared.

Table 2. Comparison of H-ART2 pattern classification model performance

pre-processing	vigi-lance	reference pattern		recogni-tion	test data (736)		average recogni-tion speed(ms)
		first	sec-ond		correct /incorrect	recog-nition rate(%)	
PF1	100	107	221	All	717/19	97.42	5.22
				Rank	717/19	97.42	4.59
				MinMax	717/19	97.42	3.25
	110	81	216	All	719/17	97.69	4.92
				Rank	719/17	97.69	4.06
				MinMax	719/17	97.69	2.76
	120	64	236	All	719/17	97.69	5.33
				Rank	719/17	97.69	3.93
				MinMax	718/18	97.55	2.68
	130	32	226	All	714/22	97.01	4.82
				Rank	714/22	97.01	2.96
				MinMax	714/22	97.01	2.16
PF2	100	115	232	All	725/11	98.51	5.05
				Rank	725/11	98.51	5.19
				MinMax	725/11	98.51	4.08
	110	84	220	All	724/12	98.37	5.17
				Rank	724/12	98.37	5.02
				MinMax	724/12	98.37	3.61
	115	42	225	All	720/16	97.83	4.74
				Rank	720/16	97.83	3.43
				MinMax	720/16	97.83	3.41
	120	26	241	All	722/14	98.10	5.47
				Rank	722/14	98.10	3.26
				MinMax	723/13	98.23	4.54
PF3	100	116	192	All	723/13	98.23	4.40
				Rank	723/13	98.23	4.51
				MinMax	723/13	98.23	3.11
	105	50	196	All	722/14	98.10	4.50
				Rank	722/14	98.10	2.99
				MinMax	722/14	98.10	2.01
	110	38	206	All	726/10	98.64	4.65
				Rank	726/10	98.64	3.37
				MinMax	726/10	98.64	2.12
	115	30	119	All	727/9	98.77	4.53
				Rank	727/9	98.77	3.01
				MinMax	727/10	98.64	2.09

b. Searching Strategies

According to searching strategies, the strategy of searching every candidate clusters(All), the strategy of searching only first clusters which have high similarity ranking with input pattern (Rank), and the strategy of searching only first clusters which have similarity ranking with input pattern by relative distance ratio between clusters (MaxMin) are experimented. First cluster selection ratio is set up to 0.3. Or, in the strategy of searching by Rank and in the strategy of searching by MaxMin, the first clusters within 30% ranks are selected.

c. Comparison Analysis

Table 2. shows the performance of facial pattern classification model(H-ART2) under various conditions. Fig. 10 and 11 show the average recognition rate and the average recognition speed per one pattern according to preprocessing methods, the number of reference patterns, and searching strategies.

Table 3. Comparison of pattern classification performance with others

class	model	conditions	recognition Rate(%)	recognition Speed(ms)
Kmeans Clustering	K-1	#C=100	95.92	2.37
	K-2	#C=150	97.15	3.52
	K-3	#C=200	97.28	4.44
	K-4	#C=250	97.83	5.51
Hierarchical Clustering	H-1	#C=100	92.39	2.30
	H-2	#C=150	94.57	3.39
	H-3	#C=200	95.52	4.45
	H-4	#C=250	97.15	5.48
Support Vector Machine	SVM-r1	RBF, C=300 → SV=655	98.78	28.07
	SVM-r2	RBF, C=500 → SV=655	98.78	28.25
	SVM-r3	RBF, C=700 → SV=655	98.78	28.17
	SVM-s1	Sigmoid, C=300 → SV=624	98.64	19.67
	SVM-s2	Sigmoid, C=500 → SV=624	98.64	18.95
	SVM-s3	Sigmoid, C=700 → SV=624	98.64	18.94
	SVM-p1	Polynomial, C=300 → SV=772	98.91	22.36
	SVM-p2	Polynomial, C=500 → SV=745	98.91	21.70
	SVM-p3	Polynomial, C=700 → SV=733	98.91	21.62
H-ART2	HART-1	ρ=100 → #C1=116, #C2=192	98.23	3.11
	HART-2	ρ=105 → #C1=50, #C2=196	98.10	2.01
	HART-3	ρ=110 → #C1=38, #C2=206	98.64	2.12
	HART-4	ρ=115 → #C1=30, #C2=199	98.64	2.09

The combined image with gray image and edge image(PF3) show better performance of recognition rate than gray image(PF1) and edge image(PF2). As the strategy of searching by relative distance ratio reduces the search space, it raises recognition speed as double without lowering the recognition rate.

B. Performance Evaluation with pattern classification model

Table 3 shows comparison of proposed pattern classification performance with Kmeans clustering, Hierarchical clustering , and SVM.

For the fairness of comparison, we make the number of reference patterns of Kmeans clustering to be equal to that of Hierarchical clustering. Three kinds of kernel functions, RBF, sigmoid, and polynomial, are applied to SVM. As the value of variable C is from 300 to 700, 600 – 700 SV(Support Vector) are generated. It shows variable C has little effect on SV.

The proposed H-ART2 model show more better recognition rate and speed than Kmeans and Hierarchical clustering. Although SVM model is very stable without the influence of parameter and shows a good recognition rate, it is difficult to control the number of SV. Therefore recognition speed becomes slow because a lot of SV are generated.

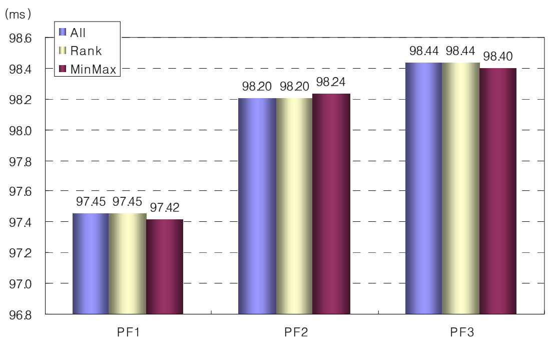

Fig. 10. Recognition rate of H-ART2

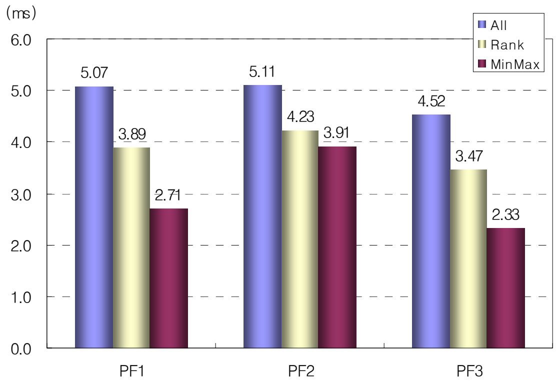

Fig. 11. Recognition speed of H-ART2

V. CONCLUSION

In this paper, we have designed the efficient classification model for face recognition which is one of the biometric techniques. First, a RGB image inputted from a frame grabber is converted into a HSV image which is similar to the human beings' vision system. Then, the coarse facial region is extracted using the hue(H) and saturation(S) components except intensity(V) component which is sensitive to the environmental illumination. Next, the fine facial region extraction process is performed by matching with the edge and gray based templates. To make an light-invariant and qualified facial image, histogram equalization and intensity compensation processing using illumination plane are performed. The finally extracted and enhanced facial images are used for training the pattern classification models.

The proposed H-ART2 pattern classification model which has the hierarchical ART2 layers make it possible to search clustered reference patterns effectively. Facial data are acquired from 10 male and female peoples at various distances and angles. Experimental results show that the proposed face recognition system is as good as the SVM model which is famous for face recognition field in recognition rate and even better in classification speed.

REFERENCES

[1] Haizhou Ai, Luhong Liang, Guangyou Xu, “Face detection based on template matching and support vector machines”, *Proceedings of Image Processing*, Vol.1, (2001) 1006-1009.

[2] R. Baldock and J. Graham, *Image Processing And Analysis*, Oxford University Press, 2000.

[3] P. N. Belhumeur, J. P. Hespanha, D. J. Kriegman, “Eigenfaces vs. Fisherfaces: recognition using class specific linear projection”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.19, No.7, (1997) 711-720.

[4] G. A. Carpenter, S. Grossberg, "ART2: Self-Organization of Stable Category Recognition Codes for Analog Input patterns", *Applied Optics*, Vol. 26, No. 23, (1987) 4919-4930.

[5] G. A. Carpenter and S. Grossberg, "The ART of Adaptive Pattern Recognition by a Self-Organizing Neural Network", *Computer*, Vol. 21, No.3, (1988) 77-88.

[6] L. Fausett, *Fundamentals of Neural Networks*, Prentice Hall, 1994.

[7] W. Gonzalez, *Digital Image Processing*, Addison Wesley, 1992.

[8] S. Haykin, *Neural Networks - A Comprehensive Foundation*, Prentice Hall, 1999.

[9] B. Heisele, P. Ho, T. Poggio, "Face recognition with support vector machines: global versus component-based approach", *8th IEEE International Conference on Computer Vision*, Vol.2, (2001) 688-694.

[10] R. Jain, R. Kasturi, and B. G. Schunck, *Machine Vision*, McGraw-Hill, ?

[11] T. Kohonen, "The Self-Organizing Map", *Proceeding of the IEEE*, Vol.78, No.9, (1990) 1464-1480.

[12] S. Romdhani, *Face Recognition using Principal Component Analysis*, The MS Thesis, University of Glasgow(UK), 1-56, April, 1997.

[13] H. A. Rowley, S. Baluja, and T. Kanade, "Rotation Invariant Neural Network-Based Face Detection", *Proceedings of Computer Vision and Pattern Recognition*, (1993) 38-44.

[14] A. Sanker and R. J. Mammone, "Neural Tree Networks", in *Neural Networks: Theory and Applications*, Academic Press, 1991.

[15] A. Sanker and R. J. Mammone, "Growing and pruning neural tree networks", *IEEE transaction Computer*, Vol.. 42, (1993) 291-299.

[16] Sergios Theodoridis, Konstantinos Koutroumbas, Pat*tern Recognition*, Academic Press, 1999.