Skin Segmentation based on Cellular Learning Automata

Ahmad Ali Abin
Computer Department
Sharif University of Technology
Tehran, Iran
abin@ce.sharif.edu

Mehran Fotouhi
Computer Department
Sharif University of Technology
Tehran, Iran
fotouhi@ce.sharif.edu

Shohreh Kasaei
Computer Department
Sharif University of Technology
Tehran, Iran
skasaei@sharif.edu

ABSTRACT

In this paper, we propose a novel algorithm that combines color and texture information of skin with cellular learning automata to segment skin-like regions in color images. First, the presence of skin colors in an image is detected, using a committee structure, to make decision from several explicit boundary skin models. Detected skin-color regions are then fed to a color texture extractor that extracts the texture features of skin regions via their color statistical properties and maps them to a skin probability map. Cellular learning automatons use this map to make decision on skin-like regions. The proposed algorithm has demonstrated true positive rate of about 83.4% and false positive rate of about 11.3% on the Compaq skin database. Experimental results show the effectiveness of the proposed algorithm.

Keywords

Skin Detection, Texture Analysis, Cellular Learning Automata.

1. INTRODUCTION

In recent years, there has been a growing research interest in the problem of segmenting skin regions in color images. Skin segmentation aims at locating skin regions in an unconstrained input image. It plays an important role in many computer vision tasks; such as face detection [1], face tracking [2], and filtering of objectionable web images [3]. Most of the existing skin detection approaches are based on the skin color (i.e., skin regions are detected by looking for pixels that have skin colors).

Cellular learning automata (CLA) is a model for systems that consist of simple elements. These simple elements improve their performance based on their neighbors' behavior and previous experiences. Nevertheless, they can expose complex behavior based on their interactions. The neighborhood properties between pixels make the cellular learning automate a good candidate for image processing.

In this paper, we propose an algorithm that combines color and texture information of skin regions with cellular learning automata to segment skin regions in color images. The presence of skin colors in the input image is first detected using a committee structure to make a decision from several explicit boundary skin models [4]. This structure maintains the benefits of different color spaces. The detected skin-color regions are then fed to a color texture feature extractor which extracts the texture features of skin regions via their statistical properties and maps them to a skin probability map. The probability map is then fed to a cellular learning automaton to make a decision on skin regions. The performance of the proposed method is evaluated using the Compaq skin database.

The organization of the paper is as follows. Skin color detection is described in Section 2. The cellular learning automata are addressed in Section 3. Proposed skin segmentation algorithm is presented in Section 4. Experimental results are shown in section 5. Finally, concluding remarks are given in Section 6.

2. SKIN DETECTION

Skin detection aims at detecting human skin pixels from a color image. The process output is a binary image that indicates the obtained skin pixels (see Fig. 1).

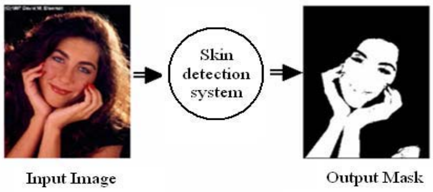


Figure. 1 A typical image and its skin detection mask.

Skin color is considered to be a useful and discriminating image feature for face and human detection, localization, and tracking. Like almost any other computer vision research fields, confounding imaging conditions (*e.g.,* change of illumination, shadows, shading, and highlights) complicate the skin detection process. In addition, the color of skin may vary among different races. Furthermore, for the same person the skin color differs significantly both in body part (*i.e.,* face versus hands) and in time (*i.e.,* after long sun exposure). A human skin color detection process is applied to decide whether a pixel belongs to a skin or to a non-skin region.

A human skin color detection process utilizes a color classifier and a color space in which colors of all objects are represented. Numerous techniques for skin color modeling and detection have been proposed. There are three broad categories of methods for

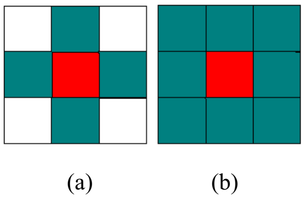
MoMM2008, November 24–26, 2008, Linz, Austria.
(c) 2008 ACM 978-1-60558-269-6/08/0011 $5.00.

skin detection and segmentation. The first category uses some explicit rules on color values [4]. In general, this type of methods is very simple to implement and is computationally inexpensive. However, it is very rigid and cannot cope with the complexity of the problem. The second category uses a non-parametric model for skin. This type of methods estimates the skin color distribution from the training data without deriving an explicit model of the skin color [5]. This category includes methods that build and use the *skin distribution map* (SMP), which is the discrete probability distribution of observed skin colors. This method is fast but requires a significant storage space. Furthermore, its performance depends heavily on the selection of the training set. The third category uses parametric models for skin color distributions. This model usually consists of a Gaussian or a mixture of Gaussian PDF and offers a more compact skin representation along with the ability to generalize and interpolate the training data [5]. In order to detect skin pixels in an image, all these methods use a number of images to build their models (or derive rules). For a particular image, the model will not coincide with the actual distribution.

All mentioned approaches are based on the same *pixel-wise* processing paradigm, in which each image pixel is individually analyzed. We believe that this paradigm should be extended; context information should be incorporated in the skin detection process. Human beings can detect skin in real scenes, or in pictures and videos without specific difficulties. However, for a human being the classification of a single pixel as skin or non-skin is a very difficult task. Humans use many high-level processes to assist the detection of skin (detection of hair, clothes, etc), and also some simple diffusion mechanisms, employed in general, for segmentation of colors and textures. When building a system that uses skin color as a feature for skin detection, the researchers usually face two main problems. First, which color space to choose and second, how exactly the skin color distribution should be modeled. In contrast, region-based methods try to take the spatial arrangement of skin pixels into account during the detection stage to enhance the performance.

3. CELLULAR LEARNING AUTOMATA

In recent years, *cellular automata* (CA) have frequently been used to model the dynamics of spatially extended physical systems. Cellular automata are a collection of cells that each adapts one of a finite number of states. Single cells change in state following a local rule that depends on the environment of the cell. The environment of a cell is usually taken to be a small number of neighboring cells. Fig. 2 shows some neighborhood options.



(a) (b)

Figure. 2 Neighborhood options: (a) Von Neumann, (b) Moore.

Learning automata (LA) are simple agents for doing simple tasks. They have a finite set of actions and at each stage choose one of them. The choice of an action depends on the state of automaton which is usually represented by an action probability vector. For each action chosen by the automaton, the environment gives a reinforcement signal with fixed unknown probability distribution, which specifies the "goodness" of the applied action. Then, upon receiving the reinforcement signal, the learning automaton updates its action probability vector by employing a learning algorithm. The interaction of the learning automaton and its environment is shown in Fig. 3.

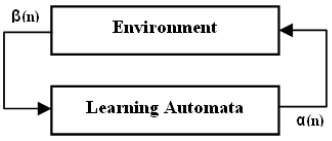


Figure. 3 Interaction of learning automata and environment.

The learning algorithm is a recurrence relation and is used to modify the action probability vector *P*. Various learning algorithms have been reported in the literature. Below, a learning algorithm, called $L_{R\text{-}\varepsilon P}$, for updating the action probability vector is given. Let α_i be the action chosen at time *k* as a sample realization from probability *P(k)* distribution and *β(k)* is the environment response to that action. In $L_{R\text{-}\varepsilon P}$ algorithm, the action probability vector is updated according to

$$\begin{pmatrix} \rho_i(n+1)=\rho_i(n)+\alpha(1-\rho_i(n)) \\ \rho_j(n+1)=(1-\alpha)\rho_j(n) \;\; \forall j \;\; j\notin i \end{pmatrix} for\ \beta(k)=0 \qquad (1)$$

$$\begin{pmatrix} \rho_i(n+1)=(1-b)\rho_i(n) \\ \rho_j(n+1)=\dfrac{b}{r-1}-b\rho_j(n) \;\; \forall j \;\; j\neq i \end{pmatrix} for\ \beta(k)=1 \qquad (2)$$

where $\rho_i(n+1)$ is the selection probability of action α_i and *a* and *b* are the decreasing and increasing factor of actions. When β(k)=0 the environment rewards the chosen action of learning automaton and when β(k)=1 the environment penalizes the chosen action. Parameter $0< a <1$ represents the step length and *r* is the number of actions for *LA* [6]. LA have been used successfully in many applications such as telephone and data network routing [7], solving NP-complete problems [8], and neural network engineering [9] to mention a few.

A CLA is a mathematical model for dynamical complex systems that consists of large number of simple learning agents [10]. A CLA is a CA in which a learning automaton will be assigned to its every cell. The learning automaton residing in each cell determines the state of the cell on the basis of its action probability vector. Like CA, there is a rule that CLA operates under it. The rule of CLA and the actions selected by the neighboring LAs of any cell determine the reinforcement signal to the LA residing in that cell. In CLA, the neighboring LAs of any cell constitute its local environment. The operation of cellular learning automata could be described as follows. At the first step, the internal state of every cell is specified. The state of every cell is determined on the basis of action probability vectors of the learning automaton residing in that cell. The initial value of this state may be chosen on the basis of past experience or at random.

In the second step, the rule of cellular automata determines the reinforcement signal to each learning automaton residing in that cell. Finally, each learning automaton updates its action probability vector on the basis of supplied reinforcement signal and the chosen action. This process continues until the desired result is obtained.

A number of applications for CLA have been developed recently; such as image processing [11] and modeling of commerce networks [12]. In [13], a mathematical methodology to study the behavior of the CLA is given and its convergence properties are investigated.

4. PROPOSED METHOD

In this paper, we propose an algorithm that combines color and texture information of skin region with cellular learning automata to segment skin regions in color images. Fig. 4 shows the overall structure of the proposed method. Each part of the proposed method is described below.

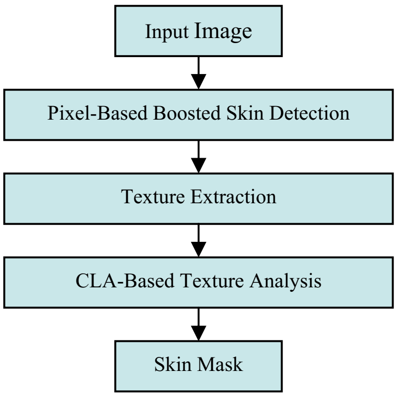

Input Image
Pixel-Based Boosted Skin Detection
Texture Extraction
CLA-Based Texture Analysis
Skin Mask

Figure. 4 Overall structure of the proposed method.

4.1 PIXEL-BASED BOOSTED SKIN DETECTION

There is no explicit evidence which shows that a special color space has the best performance for skin detection in all situations and images. Our experiments and several previous works which have been done in this area confirm this fact. Thus, if we combine different skin color spaces in an appropriate manner, we can get a model which is near to the optimum model in all situations.

In our approach, we assume a structure of classifiers that each of them is an explicit boundary skin detector in a specific color space. A boosting method called "unbiased voting" [14] is used to combine the results of the classifier and makes a better final decision. This method gives a weight to each classifier. This weight indicates the effect of each classifier result in the final decision. The conducted experiments show a direct relation between true positives (TP) and false positives (FP) rates in skin detectors. We have found that if the classifier weights are determined in an order that a balance exists among high and low TP-FP classifier weights, the best results will be achieved. The pixel-based boosted skin detector structure is shown in Fig. 5. The weight of each classifier is determined as below.

$$\omega_i = \frac{T_i}{\sum_{j=1}^{m} T_j} \tag{3}$$

where T_i is the true positive rate of each explicit skin classifier on the training data. Threshold of pixel-based boosted skin detection (ɵ) is determined empirically to detect 95.2% of all skin-associated pixels and assessment is then made in terms of the percentage of non-skin pixels incorrectly accepted. The lowest false acceptance rate found to be about 27.6%. The threshold can be change to achieve better TP versus lower FP. The aim of our proposed method is to detect the early skin region with high TP and FP and then reduce the FP by texture analysis.

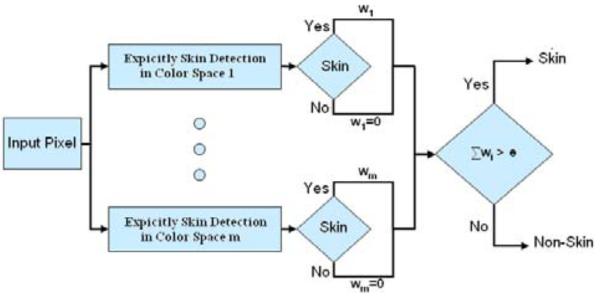

Input Pixel
Expicitly Skin Detection in Color Space 1
Expicitly Skin Detection in Color Space m
Yes
Skin
No
w1
w1=0
Yes
Skin
No
wm
wm=0
∑wi > ɵ
Yes
Skin
No
Non-Skin

Figure. 5 Structure of pixel-based boosted skin detector.

4.2 TEXTURE EXTRACTION

The texture of skin regions is somewhat smooth and if seen from far the observer might not be able to recognize any special texture for it. Pixel-based approaches judge whether a pixel is a skin pixel only by its color, while our region-based approach combines color and texture information of the skin region candidates. As a result, a non-skin region with the same color as a skin region can be discriminated by its different texture properties. Fig. 6 shows an example of two skin candidate regions with different textures.

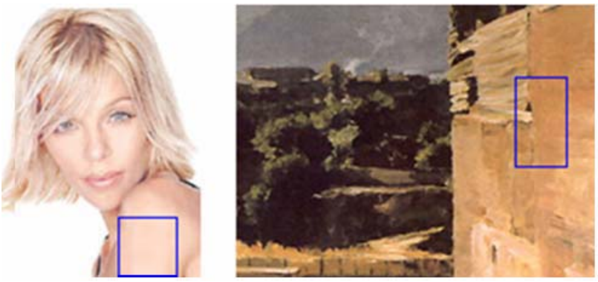


Figure. 6 Two skin candidate regions with different textures.

In this paper, we extract the texture of a region based on color information of its neighboring pixels and their Euclidian distance in the RGB color space. First, for each pixel $X_{i,j}(r,g,b)$ at location *(i,j)* of image "*I*" (with red, green, and blue values "r", "g", and "b", respectively) we calculate the Euclidian distance between the color of that pixel and all its neighbors in a W×W block around it using (4). Then, we calculate the "coarseness map" by (5).

$$D_{Y,Xi,j} = \sqrt{(r_{Xi,j} - r_Y)^2 + (g_{Xi,j} - g_Y)^2 + (b_{Xi,j} - b_Y)^2} \qquad (4)$$

$$C_{i,j} = Variance(D_{Y,X\,i,j}) \quad Y \in W \times W \; around\, X_{i,j} \qquad (5)$$

Then, we compute the "probability map" using the calculated coarseness value by (6).

$$P_{i,j} = 1 - \frac{C_{i,j}}{Max(C_{m,n})} \quad m,n \in I \qquad (6)$$

Fig. 7 shows two typical images and their related probability maps. Brighter pixels indicate greater skin probability.

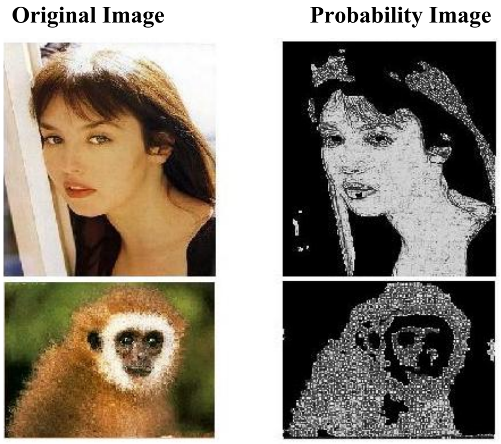


Figure. 7 Two images and their obtained probability maps.

From this figure, one can see that skin-like regions which contain a coarse texture have small and irregular probability maps while smooth skin-like regions have higher and more regular probability maps.

4.3 CLA-BASED TEXTURE ANALYSIS

In this section, we use cellular learning automata to make a decision on skin or non-skin regions using its obtained probability map. The main idea for using the cellular learning automata is to use the neighborhood relation of skin-like regions for making better decisions and thus to improve skin detection performance. CLA can propagate the skin probability of neighboring regions to all directions and then by means of this propagation can make a decision based on the overall texture information of that region.

To do so, first we create a 2-D cellular learning automata where the dimensions of probability map are calculated by the texture extraction process. Then, we allocate a dynamic structure learning automata for each cell of CLA. Each learning automaton takes two actions. These actions are related to skin and non-skin regions. The initial probability associated with the CLA comes from the probability map.

We reward or punish each learning automaton based on the selected action of the central automata and its neighbors. If the number of learning automata that select action related to skin is greater than 7, we reward the central learning automata by means of its learning algorithm. It means that if more than 7 automata report a weak texture region around it, the region around the central learning automata can be a candidate skin region. If the number of learning automat that select action related to non-skin is lower than 4, we punish the central learning automata by its learning algorithm. It means that if lower than 4 automata report a strong non-skin texture region around it, the region around the central learning automata can be a non-skin region candidate. Fig. 8 shows some states in which the central automata must be rewarded or punished.

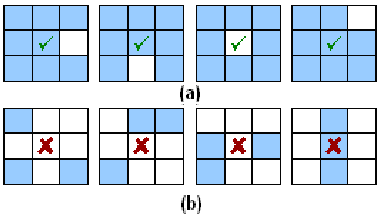


Figure. 8 Examples of (a) reward and (b) punishment states.

After several iterations, each existing automata in the CLA converges to select the skin or non-skin state with a high probability. (It means that the skin and non-skin probability associated to each learning automata converges to 0 or 1.) Thus, after several iterations, the overall system converges and the process stops. Now, we can classify each pixel as a skin or non-skin pixel by applying a threshold on its probability. Fig. 9 shows the convergence of skin probability map obtained by CLA and its thresholding.

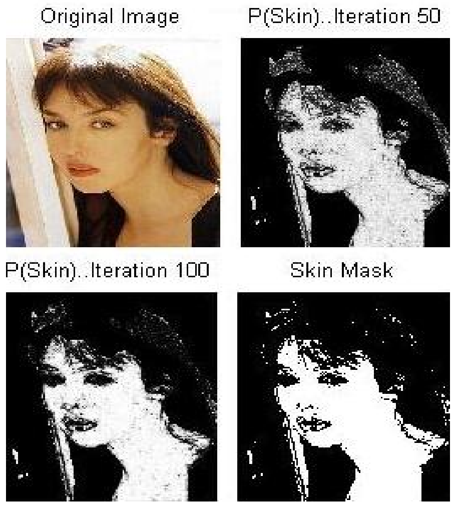


Figure. 9 Convergence of CLA for skin probability map.

5. EXPRIMENTAL RESULTS

Our skin detection system was carried out on a 2 GHz processor with 1024 MB RAM on Windows XP professional platform. We used MATLAB 7.1 and image processing toolbox 5.0.2.

For pixel-based boosted skin detection we used the color spaces RGB, YCbCr, HSV, Nrgb, and YIQ. The threshold θ was set to

0.65 to detect 95.2% of all skin-associated pixels with false positive rate 27.6%.

The texture extractor analysis system was configured with window size 3×3 and the setting of CLA was adjusted with a=0.040 and b=0.015 for iterations 30, 60, 90, 120, and 150. We tested our proposed method on the Compaq skin database (13640 web images, 4675 skin, and 8965 non-skin images) [5].

We considered the Moore neighborhood with $r=2$ for the CLA cells. The learning algorithm of each dynamic structure learning automata was considered the $L_{R-\epsilon I}$ algorithm with $a=0.040$ and $b=0.015$.

Table 1 lists the performance of the proposed methods and some previous work which were reported on this dataset in terms of TP and FP rate.

Table. 1 Performance of different skin detectors.

Detection method	Color Space	TP	FP
Bayes [15]	RGB	93.4	19.8
Elliptic Model [16]	Xyz	90	20.9
Thresholding [17]	YCbCr	82	18.7
MaxEnt. Model [18]	RGB	82.9	10
GMM (16) [5]	RGB	90	15.5
Pixel-Based Boosted	RGB, YIQ, HSV Nrgb, YCbCr	95.2	27.6
Proposed Method	RGB, YIQ, HSV Nrgb, YCbCr	83.4	11.3

As shown in Table 1, the usage of texture analysis and CLA has reduced the FP rate of the pixel-based skin detector but has also decreased the TP rate of it.

We tested the proposed method for window size 3×3 for texture extractor for iterations 30, 60, 90, 120, and 150 for CLA. The results of these configurations are shown in Fig.10. The figure shows the ROC of proposed method on Compaq skin database.

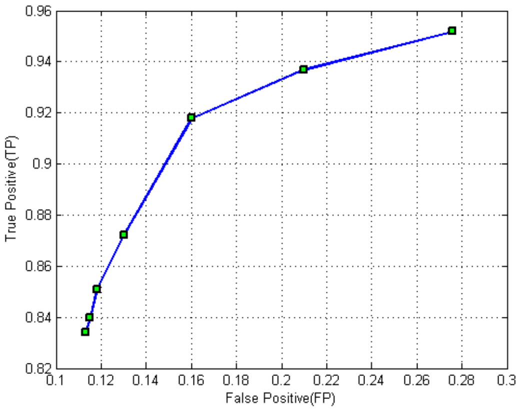


Figure. 10 ROC of proposed method.

Fig. 11 shows the result of the proposed method for some typical images of Compaq skin database. The left column shows the test image, the middle column shows the result of the proposed pixel-based boosted skin detector, and the right column shows the result of the overall system. The elapsed time of the proposed method is about 2 seconds for a 288×352 input image.

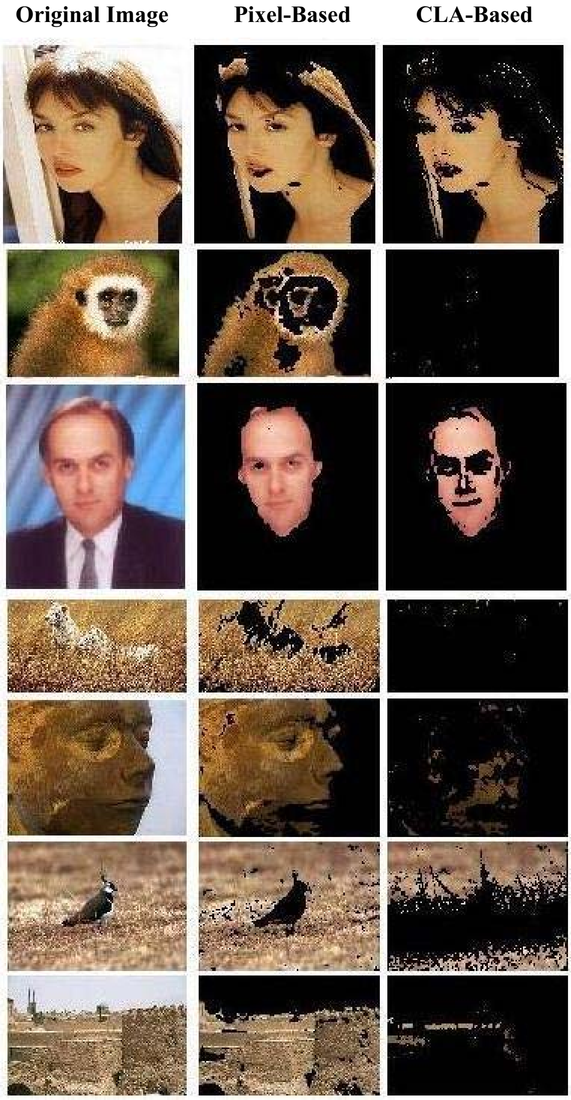


Figure. 11 Obtained skin regions.

6. CONCLUSION

The aim of this paper was to classify each query skin-like pixel as a skin or non-skin pixel not only based on its color but rather based on the existing texture around that pixel. We proposed a skin detection algorithm which extracts some raw information from color and texture of the skin-like regions. This raw information was fed to a cellular learning automaton that was able to converge to a stable state based on the related color and texture information. One can then detect the skin regions after convergence of the CLA. Our pixel-based method did not consider the texture information of the skin-like regions and made a decision only based on the color of pixels. The result of the skin detection was fed to a texture analyzer which performed based on CLA and statistical information of skin-like regions. Note that previous works achieve a good TP rate but with a high FP rate (as shown in Table 1). Our proposed skin detector obtained candidate skin regions with a TP of about 83.4% versus a FP rate of 11.3%.

One of the benefits of the proposed method compared to the conventional methods which are based on texture features is that in those methods the decision on skin regions at each block is made by using only the information of that block and its surrounding blocks, but in the proposed method the information of all image they are all counted in the final decision.

7. ACKNOWLEDGMENTS

This work was in part supported by a grant from ITRC. We also would like to thank Dr. Beigy for his help on theory and implementation of CLA.

8. REFERENCES

[1] Hsu, R.L. Abdel-Mottaleb, M. Jain, A.K. 2002. Face detection in color images. IEEE Trans. Pattern Anal. Machine Intell. Vol. 24, no. 5, pp. 696–706.

[2] Huang, F. J. Chen, T. 2000. Tracking of multiple faces for human-computer interfaces and virtual environments. Proc. IEEE Int’l Conf. Multimedia and Expo, New York.

[3] Ho, W. Watters, P. 2004. Statistical and structural approaches to filtering internet pornography. in IEEE International Conference on Systems, Man and Cybernetics, pp. 4792-4798.

[4] Gomez, G. Morales, E. 2002. Automatic feature construction and a simple rule induction algorithm for skin detection. Proc. of the ICML Workshop on Machine Learning in Computer Vision, pp. 31–38.

[5] Jones, M. Rehg, J. 1998. Statistical Colour Models with Application to Skin Colour Detection. Compaq Cambridge Research Lab, Technical Report CRL 98/11.

[6] Narendra, K. S. Thathachar, K. S. 1998. Learning Automata: An Introduction. New York, Printice-Hall.

[7] Srikantakumar, P. R. Narendra, K. S. 1982. A Learning Model for Routing in Telephone Networks," SIAM Journal of Control and Optimization, vol. 20, no. 1, pp. 34-57.

[8] Oommen, B. J. de St Croix, E. V. 1996. Graph Partitioning Using Learning Automata. IEEE Transactions on Computers. Vol. 45, no. 2, pp. 195-208.

[9] Meybodi, M. R. Beigy, H. 2002. New Learning Automata-Based Algorithms for Adaptation of Backpropagation Algorithm Parameters. International Journal of Neural Systems, vol. 12, no. 1, pp. 45-68.

[10] Meybodi, M. R. Beigy, H. Taherkhani, M. 2003. Cellular Learning Automata and its Applications. Sharif Journal of Science and Technology, vol. 19, no. 25, pp. 54-77.

[11] Meybodi, M. R. Kharazmi, M. R. 2004, Application of Cellular Learning Automata to Image Processing. Journal of Amirkabir, vol. 14, no. 56, pp. 1101- 1126.

[12] Meybodi, M. R. and Khojasteh, M. R. 2001. Application of Cellular Learning Automata in Modeling of Commerce Networks", Proceedings of 6th Annual International Computer Society of Iran Computer Conference CSICC-2001, Isfahan, Iran, pp. 284-295.

[13] Beigy, H. Meybodi, M. R. 2004. A Mathematical Framework to Study the Evolution of Cellular Learning Automata. Advances in Complex Systems, vol. 7, no. 3-4, pp. 295-319.

[14] Haykin, S. 1998. Neural Networks: A Comprehensive Foundation. Prentice Hall PTR Upper Saddle River, NJ, USA.

[15] Brand, J. Mason, J. 2000. A comparative assessment of three approaches to pixel level human skin-detection. ICPR01, pp. 1056–1059.

[16] Lee, J.Y. Yoo, S.I. 2002. An elliptical boundary model for skin color detection," Proceedings of the International Conference on Imaging Science, Systems and Technology.

[17] Phung, S.L. Bouzerdoum, A., Chai, D. 2005. Skin segmentation using color pixel classification: analysis and comparison. IEEE Trans. Pattern Anal. Mach. Intell. Vol. 27 no. 1.

[18] Jedynak, B. Zheng, H. Daoudi, M. 2003. Statistical models for skin detection. IEEE Workshop on Statistical Analysis in Computer Vision.