Advances in Environmental Biology, 8(11) Special 2014, Pages: 1216-1220

AENSI Journals

Advances in Environmental Biology

ISSN-1995-0756 EISSN-1998-1066

Journal home page: http://www.aensiweb.com/AEB/

Skin detection and isolation in the image using statistical methods

[1]Aboozar Zandvakili, [2]Babak Masoudi, [3]Habib Khodadadi

[1]Department of Computer Engineering, College of Engineering, jiroftBranch, Islamic Azad University, jiroft Iran.
[2]Department of information technology, Payamenoor university(PNU),P.O.Box, 19395-3697 Tehran,I.R of Iran.
[3]Department of Computer Engineering, College of Engineering, minabBranch, Islamic Azad University, minab Iran.

ARTICLE INFO

Article history:
Received 15 April 2014
Received in revised form 22 May 2014
Accepted 25 May 2014
Available online 15 June 2014

Keywords:
skin detection, methods based on the normal distribution, linear methods, nonlinear methods

ABSTRACT

Skin detection has different applications in computer vision and image processing. In this paper, relying on statistical methods, the issue of skin detection is investigated and given that the discussion is about color images, the precession in choosing the correct color space is an inevitable issue. HSV color space is used in this paper. Some images are used to build the data set and three features are extracted from the images and the fourth feature determines the skin and non-skin class. Another important issue is the use of an appropriate model for detecting skin regions in the image. Three groups of classification methods have been reviewed and compared. The three categories are based on normal distribution such as linear, nonlinear, and high order polynomials methods. Estimating the errors of categories are as follows. Group 1: 0.1594. Group 2: 0.00025 and group 3: 0.0776. As can be seen, non-linear methods have the best performance.

To Cite This Article: Aboozar Zandvakili, Babak Masoudi, Habib Khodadadi, Skin detection and isolation in the image using statistical methods. ***Adv. Environ. Biol., 8(11),*** *1216-1220, 2014*

INTRODUCTION

Skin detection is used in many fields such as face recognition [1], video compression, content-based image retrieval systems, filtering the images content [1,5], analysis of behavior and evaluation of face [3]. Skin detection methods are divided into two categories. the first category consists pixel-based methods in which each pixel is classified separately as skin or non-skin [3]. The second category is region-based methods. [2]. modeling techniques of skin color distribution include histogram models, Gaussian models, elliptical boundary model, artificial neural network models, maximum entropy model (Main Line Models, Hidden Markov Models, first-order model , parameter estimation model) [3]. Due to the sensitivity of skin color in images towards a variety of factors such as lighting, camera features, race (human skin color might differ from different nationalities from one continent to another continent), Personal characteristics (age, nudity, sex, body, makeup, etc), skin color detection methods are faced with problems and the performance of these methods will decrease. [2,3].

Selecting the color space is the first step and selecting a model to simulate the recognition process is the next step. In this paper, HSV color space is used because this color space is one of the perceptual color spaces and is very close to the human visual system. In [6] it was mentioned that this color space has the best performance among the other color spaces. And also three categories of parametric classification methods (linear, nonlinear, linear and high order polynomials methods) have been used. Parametric methods have a high rate of misdiagnosis of a pixel as skin pixel [1]. The aim of this article is the comparison between these categories in terms of performance to identify the skin. A question that arises here is that which category is more suitable for detecting skin? The answer is that linear methods are acceptable but non-linear methods work better.

In section 1, we introduce the color spaces and in section 2 the used methods are described. In section 3 the method of creating the data sets is mentioned and section 4 deals with the experimental results and qualitative comparison of methods and finally section 5 is devoted to conclusions.

Corresponding Author: Aboozar Zandvakili, Department of Computer Engineering, College of Engineering, jiroftBranch, Islamic Azad University, jiroft Iran.
E-mail: zandvakili.a@gmail.com

Color spaces:

Several color spaces have been proposed for skin detection. These spaces include perceptual color spaces (HSI, HSV, HSL, TSL), color-based spaces (RGB, normalizedRGB, CIE-XYZ), Vertical color spaces (YCbCr, YIQ, YUV, YES), perceptual uniform color spaces (CIE-Lab andCIE-Luv) [3]. One of the most used color spaces for storage and presentation of digital images is RGB color space. Any other color space is obtained with a linear or nonlinear transformation of RGB.

The used methods:

Methods that work based on the normal distribution:

This type of methods is parametric. Some of these methods are presented below.

Elliptical Gaussian method:

This method is a pixel-based method and its probability distribution function is as follows.

$$P(c \mid skin) = (2\pi)^{-1} \mid \sum_{S} \mid^{\frac{-1}{2}} e^{\frac{-1}{2}(c-\mu_s)^T \Sigma_s^{-1}(c-\mu_s)}$$

In relation 1, c is the color vector; μ and Σ are the distribution parameters (mean and variance). [2].

Linear method (LDC):

It has a classification that works with the minimum Mahalanobis distance. Given that the covariance matrix for each class is similar and different, it estimates the covariance and mean. The result is a linear decision boundary [7].

Independent from feature (UDC):

It has a nonlinear classification that works assuming normally distribution and uncorrelated feature.

Non-linear method (DC):

It estimates for each class of covariance and mean and the result is a non-linear decision boundary. [7].

Non-Linear methods:

Linear methods for real-world have high problems such as the high error minimum. With a smart choice for non-linear benchmark functions, we can find a way to minimize the error. Naturally, the main problem is the selection of appropriate nonlinear functions.

Decision Tree method:

Decision Tree classifies the examples by sorting them down the tree from the roots to the leaves. Each node in the tree represents a feature in the data set. The decision tree is resistant to errors. If the data set has missing values the decision tree will not face problems. [8,9]. There are some basic algorithms for decision trees such as ID3 and c4.5.

Support Vector Machine (SVM):

This model, which is abbreviated as SVM has a variety of functions such as biological research, medical image processing and etc. This model was introduced in 1995 by vapnik which is a binary classifier. It means that it has the ability to handle two-class classification at the moment.

The purpose is finding a hyperplane to separate the data into two separate classes. In general there are many hyperplanes for classification among which svm selects a hyperplane that generate greater distance between the two classes (Figure 1).

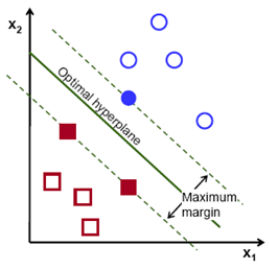

Fig. 1: a hyperplane which creates the largest margin is selected for separating. Dividing the data into two classes of skin and non skin

In many real functions, there is not a hyperplane that can linearly separate the two classes. a solution is to transfer the data from the original space into another space and then determine a hyperplane that can do a linear separation (Figure 2).

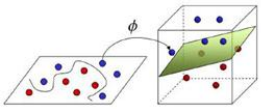

Fig. 2: Using kernel function to map into the higher dimensional space

linear and higher-order polynomials methods:

Linear classification is one of the Pattern Recognition methods in which the areas of decision making in data distribution environment are built using lines.

Logical linear method (LOGLC:)

This classifier maximizes the amount of similarity using logical function and the equivalent Perceptron network is without a hidden layer. [7].

Linear least square error method (FISHER):

It is a linear method based on least squares error and was introduced by Fisher. It pasteurizes the D-dimensional data to line m.

neighbor average method (NMC):

It calculates the mean for each class and puts the samples in the closest distance to the mean. It estimates the parameter easily. (n = number of classes, d = number of features). This method does not use the sensitive grading and previous data [7].

Neighbor mean scaled method (NMSC):

This classifier is linear and considers covariance as zero for classes and considers variance the same for all classes.

Data set:

Firstly the data set related to the RGB color space must be built and then the data sets related to HSV color space must be built.

RGB class

HSV class

$$(R,G,B) \text{ is classified as skin if :}$$
$$R > 95 \text{ and } G > 40 \text{ and } B > 20 \text{ and}$$
$$\max\{R,G,B\} - \min\{R,G,B\} > 15 \text{ and}$$
$$|R - G| > 15 \text{ and } R > G \text{ and } R > B$$

$$(H,S,V) \text{ is classified as skin if :}$$
$$H = [0,50] \text{ and } S = [0.20,0.68] \text{ and } V = [0.35,1]$$

In this paper, 556 images were used to build the data set among which 3 features are extracted and the fourth feature determines the skin or non-skin class.

The experimental results:

7.0 And 7.5 versions of MATLAB software have been used for the implementation. To investigate the usefulness of the method, several parameters have been used. True Positive parameter shows the skin pixels that were correctly identified as skin pixels. True Negative parameter shows the non-skin pixels that were correctly identified as non-skin pixels. Error parameter shows the error of used methods.

Table 1: Performance and estimation of linear methods error

True Positive	True Negative	Error	
0.9841	0.8544	0.0808	Elliptic
0.9820	0.8540	0.0820	QDC
0.9843	0.8606	0.0776	UDC
0.9893	0.6228	0.1940	LDC

Table 2: Performance and estimation of linear and high order polynomials methods error

True Positive	True Negative	Error	
0.9439	0.7374	0.1594	LOGLC
0.9885	0.6198	0.1959	FISHER
1	0.4897	0.2552	NMC
0.9981	0.5876	0.2072	NMSC

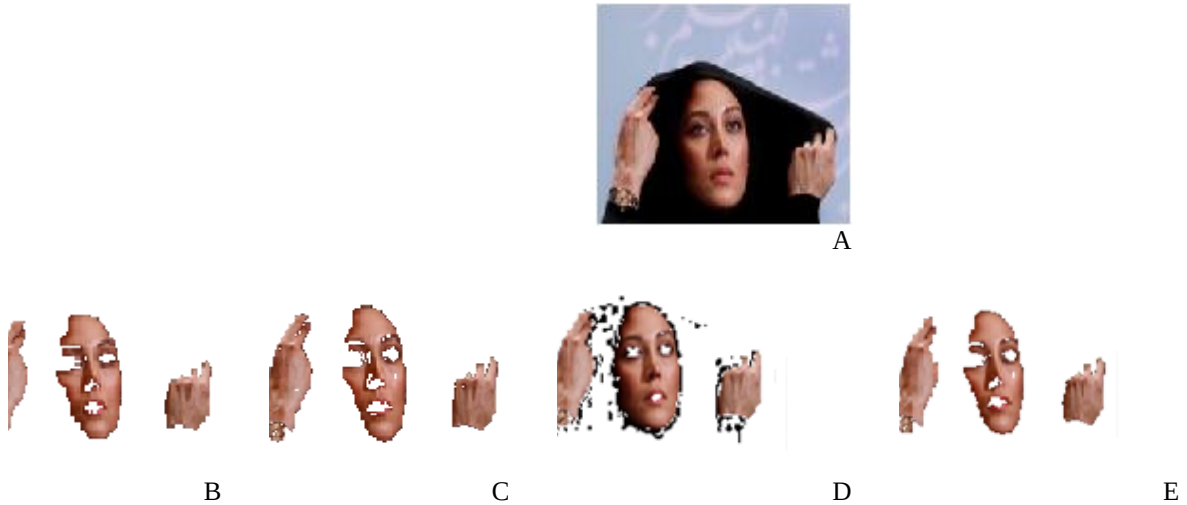

Fig. 3: A) Original image. B) QDC method. C) UDC method. D) LDC method. E) Elliptic method

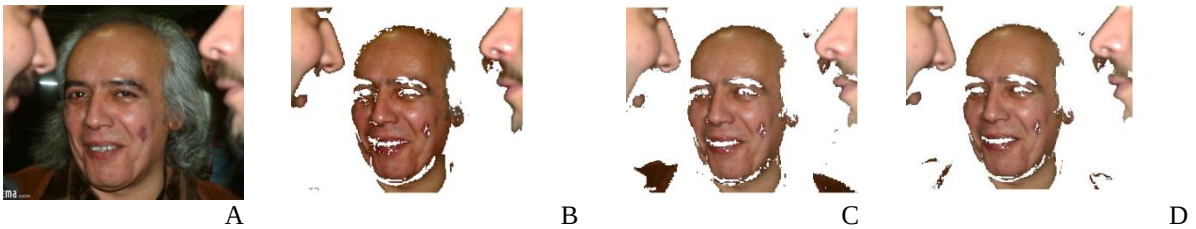

Fig. 4: A) Original image. B) DECISION TREE method. C) SVMmethod with linear core. D) SVMmethod with non-linear core

Fig. 5: A) Original image. B)LOGIC method. C) FISHER method. D) NMC method. E) NMSC method

Conclusion:

According to the results obtained in the previous section, it was found that non-linear methods have had acceptable results than the other methods with low error rate. Among the methods of this category the decision tree method is faster than the support vector machine method for training. Support vector machine approach, is

technically difficult and slow. In [11], SVM method has been used in different color spaces and the results obtained in this paper have smaller errors for the mentioned method. From the results of the first category (methods that work based on the normal distribution) it can be concluded that our data have the normal distribution in HSVdata set because their error rate is in the second place after the non-linear methods and this error is acceptable regardless of the nonlinear method. In [10] the elliptical Gaussian method in CbCr color space is used and the results are generated by higher error rate, which was expected because in [6], HSV color space had the best results. But those methods that are known as linear and high-order polynomials methods have lots of error and cannot detect the areas of skin correctly.

Pixel-based skin detection methods will be in trouble to distinguish skin-colored objects in the images because they are solely based on pixel information. To fix this problem, texture and color features can be used for skin detection.

REFERENCES

[1] Shirali-Shahreza S and M.E. Mousavi, 2008. "A New Bayesian Classifier for Skin Detection," in ICICIC '08. 3rd International Conference on Innovative Computing Information and Control, pp: 172-172.
[2] Jati. H and D.D. Dominic, 2008. "Human Skin Detection Using Defined Skin Region," in *Symposium on Information Technology,. ITSim International*, pp: 1-4.
[3] Kakumanu, P., S. Makrogiannis and N. Bourbakis, 2007. "Asurvey of skin-color modeling and detection methods," Pattern Recognition, 40: 1106-1122.
[4] Jedynak, B., H. Zheng and M. Daoudi, 2005. "Skin detection using pairwaise models," *Sience Direct*.
[5] Jedynak, B., H. Zheng and M. Daoudi, 2003. "Statistical Models for Skin Detection " in CVPRW '03. Conference on Computer Vision and Pattern Recognition Workshop, Madison, Wisconsin, USA, pp: 92-92.
[6] Zarit, B.D., B.J. Super and F.K.H. Quek, 1999. "Comparison of Five Color Models in Skin Pixel Classification," in Proceedings. International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, pp: 58-63.
[7] Werf, E.V.D., 2004. *AI techniques for the game of Go*. Netherlands: b.v., Maastricht,The Netherlands.
[8] Bishop, C.M., 2006. *Pattern Recognition and Machine Learning*. Cambridge: Springer.
[9] Mitchell, T.M., 1997. *Machine Learning*: McGraw-Hill Science/Engineering/Math.
[10] Aznaveh, M.M., H. Mirzaei, E. Roshan and M. Saraee, 2008. "A NEW AND IMPROVED SKIN DETECTION METHOD USING RGB VECTOR SPACE," *IEEE*.
[11] Xu, Z and M. Zhu, 2006. "Color-Based Skin Detection: Survey and Evaluation," in *12th Internationa Multi-Media Modelling Conference Proceedings,* Beijing, p: 10.
[12] Aznaveh, M.M., H. Mirzaei, E. Roshan and M. Saraee, 2008. "A New Color Based Method for Skin Detection Using RGB Vector Space," in *2008 Conference on Human System Interactions*, pp: 932-935.