A MARKOV RANDOM FIELD BASED SKIN DETECTION APPROACH

[1]*K. Chenaoua,* [2]*A. Bouridane*

[1]Computer Eng. Dept., KFUPM.
Dhahran 31261, Saudi Arabia
[2]School of Computer Science, Queen's University of Belfast
Belfast BT7 1NN, UK

1. ABSTRACT

A new color space is used in this paper together with a Markov Random Field model for the detection of skin pixels in colored images. The proposed color space is derived through the principal component analysis technique thus reducing the number of color components. The MRF model takes into account the spatial relations within the image that are included in the labeling process through statistical dependence among neighboring pixels. Since only two classes are considered the Ising model is used to perform the skin/nonskin classification process.

2. INTRODUCTION

Detection of human skin in color images has been the center of interest of many works in the last decade[1]. Skin segmentation is commonly used in algorithms for face detection, hand gesture analysis [2], and objectionable image filtering [1]. In these applications, the search space for objects of interest, such as faces or hands, can be reduced by detecting skin regions. Authors used different existing color spaces, combinations of different color spaces or sometimes their own proposed color space transformations [1, 3, 4]. Choice of a given color space is driven by its ability to cluster skin pixels and separate between skin and non skin pixels. Even though comparative studies have been carried out to determine a suitable color space for skin detection [5], it seems that no universal color component basis has been agreed upon.

Image segmentation is tackled as a stochastic process, and a Markov Random Field (MRF) models the joint probability distribution of image pixels in terms of local spatial interactions [6, 7]. In an MRF, spatial relationship between pixels are directly integrated and an MRF-based segmentation model can be inferred in terms of the Bayesian framework, in which various features can be used. Finally the label distribution can be obtained by maximizing the probability of the MRF model. Various MRF based segmentation models have been developed [7]. The segmentation performance of MRF segmentation models is highly dependent on the representability of the MRF parameters estimated from color and/or texture.

In this paper a new color space is proposed and an MRF model [8] is used to model the distribution of skin pixels thus taking into account the neighborhood relationship between skin pixels. In the first part of the paper a description of the newly proposed color space is given, in the second part the MRF approach is exposed and finally some results are given.

3. SKIN DETECTION

3.1 Color Spaces

Different popular color coordinates representing different classes of known color spaces [9] have been considered, thus forming a 24 component color vector. These components are: RGB, normalized RGB (rgb), YCbCr, YIQ, Lab, HSV, TSL, $(r-g)(r-b)$ and $(Y-Cb)(Y-Cr)$. The training data is taken from the ECU face detection database constructed at Edith Cowan University[10]. Each image is converted from the RGB color space, to different color spaces. The color components are considered as features. A feature vector would then be made up of the following 24 features (Eq. 1).

$$X = \begin{bmatrix} R & G & B & rn & gn & bn & Y_{cbcr} & Cb & \ldots \\ Cr & Y & I & Q & H & S & V & L & \ldots \\ a & b & T & S_{ts} & rg & rb & YCB & YCR \end{bmatrix} \quad (1)$$

3.2 Color Components Reduction

The dimensionality of the color vector is transformed to a reduced set of linearly transformed features, that is representative of the original feature set. The above vector (Eq. 1) contains all occurrences of skin pixels in the 24 different components. The number N of data pixels is over 13 million pixels, giving a raw data vector of $24 \times N$ components. If C is the covariance of X and $\bar{X}$ is the vector mean, X is linearly transformed to a lower dimensional vector Y of dimension k such that $(k < n)$ using the transformation (Eq. 2):

$$Y = A_k^T (X - \bar{X}) \quad (2)$$

where A is an $n \times k$ matrix whose columns are the k orthogonal eigenvectors corresponding to the largest k eigenvalues of the covariance matrix C.

The covariance matrix is diagonalized using the transformation:

$$\Sigma = A_k \cdot \Lambda \cdot A_k^T \quad (3)$$

where Λ is the diagonal matrix whose diagonal elements are the eigenvalues of C: $\lambda_1 > \lambda_2 > \ldots > \lambda_n$. The new vector $Y = A_k^T \cdot X$ will be a $k \times N$ element vector, where all the row elements are uncorrelated.

3.3 Common Chrominance Components

Different combinations have been proposed for the choice of the color components. When only the most common chrominance components are used, acceptable results are obtained. This combination reduces the number of components from
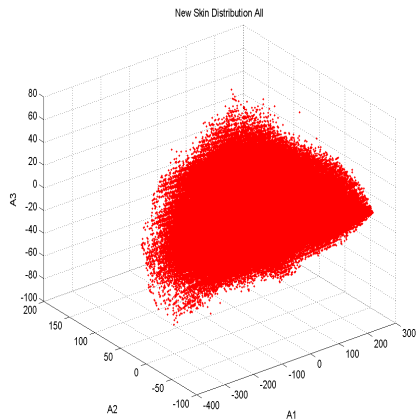
Figure 1: Skin Cluster in the New Color Space

23 to 3. A plot of the Receiver Operating Characteristic (ROC) is used to show the performance of classifiers. The main parameter derived from the ROC plot is the area under the curve (AUC), that is used in our case as a performance parameter, to compare our algorithm with work in the same field [11]. When used with an elliptical classifier [4, 12], the new approach gives a good AUC compared to [11] (Table 1).

The above combination of color components has been used for the detection of human skin in color images and shows to give good clustering of color pixels (Figure 1) and good AUC (Table 1). The next part is dedicated to the MRF model and its use for skin detection in this application.

3.4 The Elliptical Model

If X is the pixel color value, the elliptical boundary model is defined as:

$$\Phi(X) = (X - \Psi)^T \Lambda^{-1} (X - \Psi) \tag{4}$$

The model parameters (Ψ and Λ) are estimated by:

$$\Psi = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{5}$$

$$\Lambda = \frac{1}{n} \sum_{i=1}^{n} f_i (X_i - \bar{X})(X_i - \bar{X})^T \tag{6}$$

where n is the total number of skin pixels, f_i the frequency of occurrence of color i, and n is the total number of distinctive training color vectors. Finally a pixel with color X is classified as skin if:

$$\Phi(X) \leq \Theta \tag{7}$$

where Θ is a threshold value. Assuming the distribution of skin pixels fits an ellipsoid, the threshold value Θ is determined using the approach presented in [13] and initially used by [14] for face detection.

4. THE MRF MODEL

4.1 Overview

Markov Random Fields (MRF)[8], introduced for the first time in image analysis by Geman and Geman [6], provide a

good tool for modeling vision problems within the Bayesian framework using spatial continuity.

An MRF is a stochastic process in which spatial relations within the image are included in the labeling process through statistical dependence among neighboring pixels. A standard MRF model consists of two components: a region labeling component and a feature modeling component. The region labeling component imposes a homogeneity constraint on the image segmentation process, while the feature modeling component functions to fit the feature data. Color and texture are often used as features for the segmentation of colored images. A constant weighting parameter is used to combine the two components. This model works appropriately in a supervised environment; however, in an unsupervised environment, the model does not work consistently. The proposed implementation scheme, similar to [15], combines the two components using a variable weighting between them to allow the MRF model work in an unsupervised manner. As the number of classes is limited to two classes only, viz. skin and non-skin the model is considered as an Ising model [8].

4.2 MRF in Image Segmentation

The image pixels are indexed by a rectangular lattice S in which each pixel s is characterized by the gray level y_s from the set $y = y_s : s \in S$. The labeling process consists of assigning a label to each pixel $s \in S$ with a class labeling representing the pattern class in the image. A label set is defined by: $L = \{1, \ldots, C\}$, where C is the total number of label classes in the image. A labeling is denoted by $x = \{x_s : x_s \in L, s \in S\}$, where $x_s = l$ indicates that the class label S is assigned to pixel s. The goal is to find the labeling $\hat{x}$ of the image, which is the estimation of the true unknown labeling x.

$$\hat{x} = arg_x max\{P(x/y)\} \tag{8}$$

Assume $Y = y$ a feature vector extracted from the image. Therefore, according to Bayes theorem:

$$P(x/y) = \frac{P(y/x)P(x)}{P(y)} \tag{9}$$

where $P(y/x)$ is the conditional probability density function of the image y and $P(x)$ is the prior density of the labeling x. The prior probability of the image $P(y)$ is independent of the labeling x and hence can be disregarded, therefore:

$$\hat{x} = \arg_x max\{P(y/x)P(x)\} \tag{10}$$

4.3 MRF Model Property

Let $X = X_s, s \in S$ denote a family of random variables indexed by site s, Ω denote the space for all possible configurations of X. X_s assumes the labels of segments from a finite set of labels at location $s \in S$ over an $M \times N$ lattice. Let $\mathcal{N}_s$ be a general neighborhood system of S. An nth order neighborhood $\mathcal{N}_s$ of s is defined as: $\mathcal{N}_s^n = \{r | d(s, r) \leq n, r \neq s\}$ where, $d(\cdot)$ is a distance function. X is an MRF with respect to the neighborhood system $\mathcal{N}_s$ if:

$$P(X = x) \geq 0 \quad \forall x \in \Omega \tag{11}$$

$$P(X_s = x_s \mid X_r = x_r, r \neq s, \forall r \in S) = P(X_s = x_s \mid X_r = x_r, r \neq s, \forall r \in \mathcal{N}_s) \tag{12}$$

where $P(\cdot)$ and $P(\cdot|\cdot)$ denote the probabilities and conditional probabilities, respectively. The expression above (Eq.12) states that the value X_s at location S only depend on its neighbors.

4.4 The Hammersley-Clifford Theorem

An MRF can equivalently be characterized by a Gibbs distribution given by:

$$P(X) = \frac{1}{Z} \cdot \exp\left\{ -\frac{U(X)}{kT} \right\} \quad (13)$$

T is the temperature, k Boltzmann's constant and Z is a normalization constant given by: $Z = \sum_{c \in C} exp\left\{ -\frac{U(c)}{kT} \right\}$, U is the energy function, given by the sum of clique potentials V_c over all possible cliques $c \in C$ (Eq.14). A Clique $c \in C$ is defined as a subset of S such that every pair of distinct sites in c are neighbors. A single site s is also defined as a clique.

$$U(X) = \sum_{c \in C} V_c(X) \quad (14)$$

The value of U depends on the local configuration of the clique c. The Gaussian distribution is a special member of the Gibbs distribution family.

5. THE MRF BASED ALGORITHM

The color space discussed previously is used as the basis for the skin classification process. The MRF model discussed herein is used to model the distribution of skin pixels.

5.1 The Ising Model

This model is initially proposed by Ising to explain the magnetic behavior of ferromagnetic materials. The Ising model considers an idealized system of interacting particles, arranged into a regular planar grid. Each particle can have one of two magnetic spin orientations, generally labeled as: up $(+1)$ and down (-1). Each particle interacts only with its neighbors; the contribution of each particle to the total energy of the system depends upon the orientation of its spin x compared to it neighbors. Adjacent particles that have the same spin are in a lower energy state than those with antithetic spins. Given the spin orientations of all particles in the system, the total energy may be computed (Eq. 15).

$$E_R = -J \sum_{x_j \in \mathcal{N}_i} \delta(x_i, x_j) \quad (15)$$

The parameter J is the Ising parameter, δ the Kronecker symbol and $\mathcal{N}_i$ is the neighborhood of x_i. The parameter J determines the strength of the spatial interaction between particles or pixels in our case. When J is positive, agreement between neighboring pixels x_i and x_j decreases the energy. Low energy configurations are more probable: at low temperatures, the energy of a configuration is very important in determining its likelihood, therefore the most likely states are those with lowest energy. At high temperatures, energy is less important, and hence states with high entropy are not unlikely.

5.2 The Elliptical-Model Based MRF

Energy due to the first component $P(y/x)$ in the conditional labeling (Eq. 10) is equivalent to the Elliptical model (Eq. 6), and hence can simplify computations in the proposed algorithm (Algorithm 1). Assuming a gaussian distribution of skin pixels, the formula in (Eq. 4) is used to evaluate the energy term in (Eq. 13).

5.3 The Compound Model Based MRF

Assuming that the set C represents all possible configurations, and $E(c)$ is the energy of configuration $c \in C$. Statistical mechanics states that the probability of any configuration c is:

$$P(c \in C) = \frac{1}{Z} exp\left\{ -\frac{E(c)}{kT} \right\} \quad (16)$$

Finally the total energy E is found as the sum of the two components already mentioned (Eqs. 15,16):

$$E_T = E_R + \alpha E_F \quad (17)$$

Considering equations (17) and (16) the probability of a combined state given the two models is given by:

$$P(c \in C) = \frac{1}{Z} exp\{-E_T\} \quad (18)$$

The term kT can be omitted, thus simplifying computations and so is the constant Z, which finally yields the following to compute the probability $P(c \in C)$:

$$P(c \in C) \propto exp\{-E_T\} \quad (19)$$

5.4 The Annealing Scheme

Maximizing the probability in (Eq. 10) is in fact equivalent to minimizing the energy function (Eq. 17) made up of the sum of the two components (Eq. 15 and 16). Such minimization is done through the maximum a posteriori criterion (MAP) [6]. Although mathematically simple, this type of MAP estimation clearly presents a computationally infeasible problem. Therefore, optimal solutions are usually computed using some iterative optimization, or minimization, techniques. In this paper, the Gibbs/Metropolis sampler [6] is adopted. The annealing scheme used is the fast logarithmic scheme[6], where cooling is performed at each iteration ($Niter$) as follows:

$$T = \frac{1}{log(Niter+1)} \quad (20)$$

The annealing process is accelerated by eliminating, at each image scan, small holes generated by the algorithm. Patches that are less than 1% of the largest skin area are removed at this stage.

5.5 The Proposed Algorithm

In the algorithm (Algorithm 1), the aim of using the parameter α (Eq. 17) is to have the effect of the different fields at different temperatures. At high temperatures, α is high, therefore the effect of E_F (Algorithm 1) is dominant. However, at low temperatures, α is small, and therefore it is the field E_R that dominates. γ, c_1 and c_2 are constants. Experimentally, the following values have been adopted from the work by Huawu et al. [15], where it was determined that: $\gamma = 0.9$, $c_1 = 80$ and $c_2 = 1/K$ (where K is the dimension of

Algorithm 1 : Proposed Algorithm

1. Convert Image to Common Chrominance Component Space
2. Perform skin detection using the elliptical model [4]
3. $Niter = 1$; T =initial temperature
4. For each pixel x_i visited only once per every image scan, compute:

- $E_R = -J \sum_{x_j \in \mathcal{N}_i} \delta(x_i, x_j)$

- $E_F = (X - \Psi)^T \Lambda^{-1} (X - \Psi) - \frac{3}{2} \log(2\alpha) - \frac{1}{2} \log |\Lambda|$

- $\alpha = c_1 \cdot \gamma^{Niter} + c_2$

- $E_T = E_R + \alpha \cdot E_F$

- $P_b = exp\{-(E_T)\}$

- Generate a random number $U(p)$ If $P_b < U(p)$ then change the pixel class-label

5. Update parameters:

- $Niter = Niter + 1$

- $T = \frac{1}{log(Niter+1)}$

6. Remove regions whose areas are smaller than 1% of the largest skin region
7. Repeat steps 2 to 6 until, either the temperature has reached a very small value, or the difference in number of detected skin pixels between two successive results is less than 1 *percent* of the total size of the image.

the feature space) are appropriate values for a variety of images. The Ising parameter J (Eq. 15) is taken equal to 0.85. These values are used for all results presented in this paper.

6. EXPERIMENTAL RESULTS

The approach when tested on different images (Fig. 2, 3), has significantly improved the elliptical model [4]. The AUC is improved 0.925 compared to the 0.904 in case of the common chrominance components (Table 1). Notice that the ROC for the proposed MRF model is less than the ROC for the All component model at high false alarm values. This is mainly due to the fact that the range of threshold values (Eq. 7) used for the present model does not span a wider range as for the other models. However, the MRF based approach still gives a better AUC. However, in some cases (Fig. 5), due to

Components	AUC
Chrominance and Luminance Using MRF	0.925
Chrominance and Luminance only[4]	0.904
Common Chrominance[4]	0.897
Chrominance[4]	0.887
Reference Paper[11]	0.852

Table 1: Area Under Curve For Different Techniques

the fact that the background spreads over a large area and has a color similar to that of human skin, the algorithm failed to separate most human skin pixels from the background. Even though the number of iterations has been increased to over 21 iterations. However, the background patches that are similar to skin have been reduced. The algorithm can further be improved by introducing a texture feature assuming that skin regions are mostly characterized by a smooth texture as can be seen in (Fig. 6), where regions that have skin like color but with rough texture have been classified as skin.
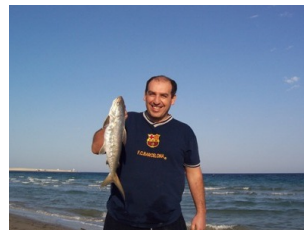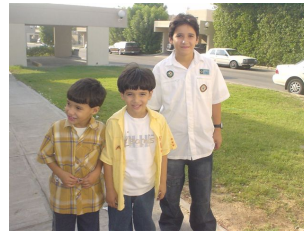

(a)
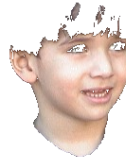
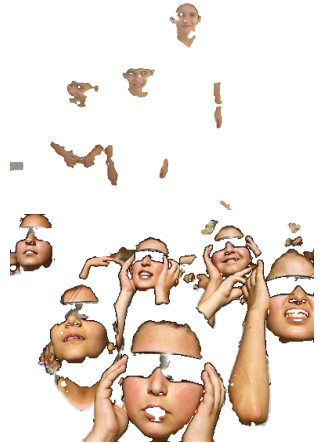(b)   (c)

Figure 2: Successful Case: (a) Original (b) Initial Skin Segmentation and (c) Final Skin Images after 7 iterations

(a)   (b)

Figure 3: Successful Cases: (a)Original and (b)Final Skin Images

7. CONCLUSION

A new color space together with an elliptical model used as a skin classifier are proposed. An MRF model is used to model the skin neighborhood interrelation. The use of the new color space combined with an MRF model has improved the skin detection process. However, in cases where the image contains objects having skin like color objects that spread over large areas within the scene, the algorithm failed to detect human skin. This can be improved
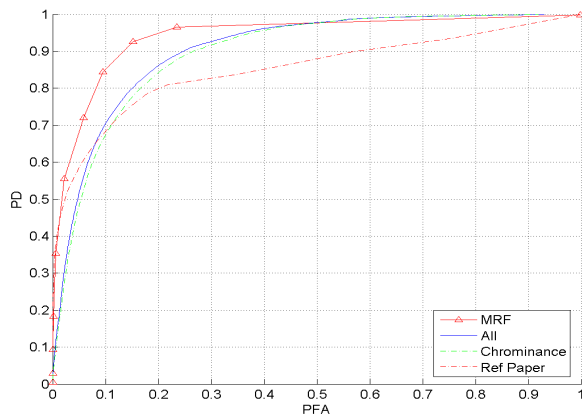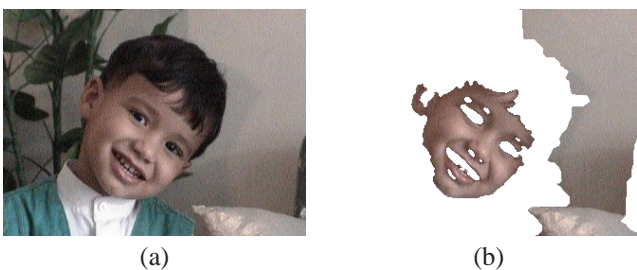
Figure 4: ROC Curves

(a)　　　　　　　　　(b)

Figure 5: Image with Skin Like Background: (a)Original and (b)Final Skin Images after 21 iterations

(a)　　　　　　　　　(b)

Figure 6: Image with Texture: (a)Original and (b)Final Skin Images after 21 iterations

by introducing a texture feature in the algorithm, which can be done if further investigations of human skin appearances and textures are carried out.

REFERENCES

[1] Vezhnevets V. Sazonov V. Andreeva A., "A survey on pixel-based skin color detection techniques," in *Graphicon*, 2003.

[2] Ming Hsuan Yang and Narendra Ahuja, "Face detection and gesture recognition for human-computer interaction," *Kluwer Academic Publishers*, 2001.

[3] G. Sanchez M. Gomez and Sucar L.E., "On selecting an appropriate color space for skin detection," Proceedings of the Second Mexican International Conference on Artificial Intelligence, 2000, vol. 2313, pp. 69–78.

[4] Chenaoua K. Bouridane A., "PCA based choice of representative colors for skin detection," EUSIPCO, 2005.

[5] Shirazi M. N. Fukamachi H. Terrillon, J.-C. and S. Akamatsu, "Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images," in *Proc. of the International Conference on Face and Gesture Recognition*, 2000, pp. 54–61.

[6] Geman D. Geman, S., "Stochastic relaxation, gibbs distributions and the bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell*, pp. 721–741, 1984.

[7] S. Z. Li., "Markov random field modeling in computer vision," *Springer-Verlag*, 2001.

[8] Gerhard Winkler, "Image analysis, random fields and markov chain monte carlo methods a mathematical introduction," *Springer-Verlag*, 2000.

[9] N. Plataniotis A.N. Venetsanopoulos., "Color image processing and applications," *Springer-Verlag*, 2000.

[10] D. Phung S. L. Bouzerdoum A. Chai., "Skin segmentation using color and edge information," in *ISPA*, 2003, pp. 525–528.

[11] Jones M. J. Rehg J. M., "Statistical color models with application to skin detection," in *Proc. of the CVPR 99*, 1999, vol. 1, pp. 274–280.

[12] Jae Y. Lee and Suk I. Yoo, "An elliptical boundary model for skin color detection," in *Proc. of the 2002 International Conference on Imaging Science, Systems, and Technology*, 2002.

[13] Chenaoua K. Bouridane A., "A new approach for the choice of a color for skin detection," Irish Machine Vision and Image Processing Conference, August 2005.

[14] D. Phung S. L. Bouzerdoum A. Chai., "A novel color model in YCBCR color space and its application to human face detection," in *ICIP*, 2002, pp. 289–292.

[15] D.A. Huawu Deng; Clausi, "Unsupervised image segmentation using a simple mrf model with a new implementation scheme," in *Proceedings of the 17th International Conference on Pattern Recognition*. ICPR, 2004, vol. 2, pp. 691–694.