Skin-color detection using fuzzy clustering

Youssef Chahir and Abder Elmoataz

Abstract— **In this paper, we present the application of the fuzzy c-means clustering algorithm to the skin-color segmentation problem. We address the problem of identifying skin-color and we adapt a spatial data mining method to this task and integrate with a segmentation method to identify significant skin-color regions in an image. The proposed algorithm is able to take into account both the distributions of data in hybrid space and the spatial interactions between neighboring pixels during clustering. Evaluation of the results indicates that the method can be used as an automatic tool for mining skin images.**

I. Introduction

Skin-color modeling is a crucial task for several applications of computer vision. Problems such as face detection in video are more likely to be solved if an efficient skin-color model is constructed. Most potential applications of skin-color model require robustness to significant variations in races, differing lighting conditions, textures and other factors.

Given the fact that a skin surface reflects the light in a different way as compared to other surfaces, we relied once again, on data-mining techniques to define a skin color model which enables the classification of image pixels into skin ones or non skin ones[1].

An important step in the image classification process is color segmentation of the image into homogeneous skin-color regions and non-skin color regions in color space that is relatively invariant to minor illuminant changes. The segmentation is used to localize and identify homogeneous regions in a picture by perceptual attributes which include the size, the shape and the texture and/or color information. This operation is necessary for any manipulation of analysis of image by computer, understanding and interpretation. After segmenting out skin from images, this can be useful for identifying faces, hand sign recognition, offensive content such pornography.

There are several strategies for segmenting images; their performances depend largely on the type of images to be processed and on a priori knowledge relative to the object features. These methods can be roughly classified into several categories: contour-based methods [2], region-based methods[3], variational methods such as the global optimization approach minimizing an energy function or some Bayesian criteria [4], active contours models[5].

a number of rules) the boundaries skin cluster in some colorspace.

Peer et al. present a simple skin classifier through a number of rules [6]. This method have attracted many researchers [7], [8], [9]. The main difficulty achieving high recognition rates with this method is the need to find both good colorspace and adequate decision rules empirically. In other hand, there have been proposed a method that uses machine learning algorithms to find both suitable colorspace and a simple decision rule that achieve high recognition rates [10].

Several face detection and tracking algorithms [11,12,13,14] use a histogram based approach to skin pixels segmentation. The colorspace (usually, the chrominance plane only) is quantized into a number of bins. These bins, forming a 2D or 3D histogram are referred to as the lookup table (LUT).

$P(c|skin)$ and $P(c|\neg skin)$ are directly computed from skin and non-skin color histograms. The prior probabilities $P(skin)$ and $P(\neg skin)$ can also be estimated from the overall number of skin and non-skin samples in the training set [11,15,16]

The paper is organized as follows. Section 2 is devoted to description of our approach used for skin detection. In sections 3 we discuss and present some results. In Section 4 the conclusion is drawn.

II. Combining decision rules and fuzzy clustering

A. *Data-mining based learning*

Let the set of pixels Ω be extracted and pre-processed automatically from training images and corresponding binary masks. We thus have a two classes classification problem, each pixel ϖ being associated with its label $C(\varpi)$: skin-color or non skin-color.

In our work, we computed for each pixel its representation in various normalized color spaces: RGB, HSV, YIQ, YCbCr, CMY in order to find the most discriminative set of color axes. This leads to a feature vector composed of 14 exogenous variables as illustrated by fig.1-a. An excerpt of the learning dataset is illustrated by Fig.1-b. Associated with each pixel feature vector is the

Manuscript received November 15, 2005

Y. Chahir is with the Department of Computer and Information Science at University of Caen, in the GREYC - URA CNRS 6072 laboratory, 14032 Caen Cedex , FRANCE, (phone: 33/0231567375, fax: 33/0231567330, e-mail: chahir@ info.unicaen.fr).

A. Elmoataz, was with Cherbourg University , LUSAC lab, Cherbourg, FRANCE (e-mail: abder@greyc.ensicaen.fr).

class label C : 1 for skin color and 0 for non skin color [1].

r	Normalized R
g	Normalized G
b	Normalized B
H	Hue
S	Saturation
V	Value
Y	Luminance
I	Inphase
Q	Quadrature
Cr	Chrominance
Cb	Chrominance
C	Cyan
M	Magenta
Y	Yellow

a – colour features

P	V1	V2	V3	V4	V5	V6		C
1	0.061055	0.031818	0.155225	0.071557	0.056528	0.071690	…	1
2	0.071105	0.172834	0.192698	0.082228	0.167058	0.082224	…	1
3	0.049195	0.022229	0.048231	0.054901	0.028214	0.055375	…	1
4	0.049195	0.022229	0.048231	0.054901	0.028214	0.055375	…	1
5	0.071105	0.172834	0.192698	0.082228	0.167058	0.082224	…	0
6	0.054431	0.022827	0.046170	0.061891	0.027777	0.062345	…	0
7	0.049195	0.020797	0.024770	0.054901	0.028214	0.055375	…	1
8	0.049195	0.028440	0.038839	0.054901	0.028214	0.055375	…	1
9	0.071105	0.172834	0.192698	0.082228	0.167058	0.082224	…	0
10	0.129661	0.421915	0.329804	0.339445	0.383743	0.30221	…	1

b – excerpt of learning dataset

Figure 1 – color features (a – left) and excerpt of learning dataset (b – right)

We used SIPINA [5] technique for training. As result, a hierarchical structure of classification rules of the type "IF...THEN..." is created (Figure 2).

In preliminary experiments we found that the hybrid space composed of HSV color space and spectral distribution is much better than normalized (r,g,b) for estimation and prediction of skin color distribution in image sequences [18]. We want to apply what we have learned to create a general image segmentation method based on modified c–mean clustering. We define a multidimensional image with four channels H, S, V and D. HSV is obtained by conversion of RGB, while D represent the spectral distribution of pixels in the image.

The pixel distribution on each channel is represented by an interval of approximated means (and standard deviations) $\mu_{Channel}$ ($\sigma_{Channel}$) (cf. Figure 2: graph of induction)

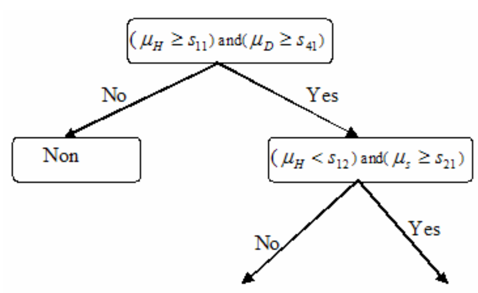

Figure2: Example of decision rules

Our objective is to group the pixels in to clusters based on their similarity based on perceptual attributes (such the skin color) and the neighborhood to force the algorithm to create coherent regions.

The image is firstly pre-processed in order to eliminate noise of the original, improving the signal-to-noise ration (SNR) of a given image. This is an essential step in image analysis especially in high noise situations that can abrupt the shape information. Image denoising has two principals requirement. One is to smooth all homogeneous regions that contain noise and the other is to retain in an accurate way the location of the boundaries that defines the shape of the represented objects. To preserve this, the original image is blurred with a Gaussian blur (Gaussian filtering).

B. Fuzzy C-means Approach:

Fuzzy C-means Clustering (FCM), is an clustering technique which employs fuzzy partitioning, in an iterative algorithm. The aim of FCM is to find cluster centroids that minimize a dissimilarity function. The Fuzzy C-Means algorithm (FCM) [19] minimizes (JA) given by:

$$J_A = \sum_{i=1}^{C} \sum_{j=1}^{N} (u_{ij})^m (d_{ij})^2$$

with the constraint $\sum_{i=1}^{C} u_{ij} = 1 \quad \forall \ j \in [1, N]$ and $u_{ij} \in [0,1]$

u_{ij} is the is the membership value at pixel j in the class I.

U= [u_{ij}] is a C x N matrix called the constrained fuzzy C-partition matrix. m ϵ [1,∞] is a weighting exponent, known as the fuzzifier. c_i is the centroid of cluster i; d_{ij} is the Euclidian distance between i_{th} centroid(c_i) and j_{th} data point. N is the total number of pixels in image and C the number of clusters to be found.

It is usually used to cluster a set of data based on their respective relating distances. The parameter m is a weighting exponent on each fuzzy membership and determines the amount of fuzziness of the resulting classification. The FCM objective function is minimized when high membership values are assigned to pixels whose colors are close to the centroid of its particular class, and low membership values are assigned when the pixel data is far from the centroid [20]

H. Frigui, and R. Krishnapuram [21], propose a Competitive Agglomeration Clustering (CA) schema where they add a second regularization term to prevent overfitting the data set with too many prototypes. The resulting objective function J_B is defined as:

$$J_B = \sum_{i=1}^{C} \sum_{j=1}^{N} (u_{ij})^2 (d_{ij})^2 - \alpha \sum_{i=1}^{C} \left[\sum_{i=1}^{C} u_{ij} \right]^2$$

This objective function combines two components. The first one is similar to the FCM (when α =0). The effect of the neighbors' term is controlled by the parameter α. The relative importance of the regularizing term is inversely proportional to the signal-to-noise ratio (SNR). Lower SNR would require a higher value of the parameter α.

We apply a modified fuzzy c-mean algorithm (MFCM) proposed yet in [22] to classify all pixels in a given image into C classes by minimizing the following objective function:

$$J = \sum_{i=1}^{C} \sum_{j=1}^{N} (u_{ij})^2 d^2(x_j, c_i) - \alpha \sum_{i=1}^{C} p_i \log(p_i) \quad (1)$$

With $p_i = \frac{1}{N} \sum_{j=1}^{N} u_{ij} = \frac{N_i}{N}$

N_i is the cardinality of cluster i. p_i is interpreted as the probability of cluster i. $d_{ij}^2 = d^2(x_j, c_i)$ is the standard Euclidian distance between ith centroid (ci) and jth pixel.

In our case, in order to take account of the distribution in the four channels we choose a distance function as $d_{ij} = \sum_{k=1}^{4} \omega_k (x_{jk}, c_{ik})$ and probability as $p_i = \sum_{k=1}^{4} \frac{N_{ik}}{N}$

ω_k serve as weighting constants between the different channels. c_{ik} and x_{jk} represent respectively centroid of the class i in the channel k and jth pixel in the same channel.

N_{ik} is interpreted as the cardinality of cluster i in the channel k.

$$u_{ij} = \frac{1/d_{ij}^2}{\sum_{k}^{C} (1/d_{kj}^2)} + \frac{\alpha}{2Nd_{ij}^2} \left[\log p_i - \frac{\sum_{k}^{C} (\frac{1}{d_{kj}^2} \log p_k)}{\sum_{k}^{C} (1/d_{jk}^2)} \right]$$

After updating centroïds and memberships, a priori probabilities of clusters are updated. We discard clusters whose probabilities are below a given threshold ε . ε should be small enough to only discard non-significant clusters. We assign each pixel to the cluster in which the membership value is the highest. Detailed algorithm of the modified fuzzy c-means is described in the following steps [23] (cf. table 1).

In our algorithm the number of classes, C is determined automatically by choosing a high value of C and eliminating the class i with the smallest probability p_i. This is the main difference with the classical algorithm for which the number of cluster is fixed.

III. Results

The figure 3 shows the results of pixels classification pixels according to general FCM method and our modified FCM approach. As we can see, in the general fuzzy clustering method, the results depend on the number of clusters. The clusters are not optimal and the regions are not usually homogeneous. When we combine our skin model and MFCM clustering, the areas are more homogeneous and more compact.

TABLE I : Algorithm: MFCM

Step1	
Step 1	Randomly initialize the membership matrix U
Step 2	Update the centroids(ci) for each cluster
Step 3	Update the a priori probabilities p_i of each cluster, if $p_i <= \varepsilon$, discard cluster i
Step 4	Compute the Euclidian distance d_{ij} between centroids and pixels
Step 5	If $\left\| J^n - J^{n+1} \right\| > \xi$ then $n = n+1$, update α Update the membership matrix (U) . Go to step 2

These results are interesting insofar as they take account at the same time our skin model, which is composed of HSV components and spectral distribution, and the neighbourhood of pixels. Nevertheless, there remains small areas which it will be necessary to eliminate (noises) according to the applications. It will also be necessary to supplement this phase of classification by a phase of segmentation which must preserve contours of the representative homogeneous areas and also allow the space localization of the coherent areas. Let us recall, that an homogeneous area is a compact area which is made up of close pixels (near to the pixels of skin). A homogeneous area is considered coherent if it represents beyond a certain threshold (% of the image).

IV. Conclusion

We applied the FCM method to perform skin detection based on decision rules in hybrid space. The results are satisfactory and will be extended to accommodate as part of web filtering system.

In this paper, we have presented a skin-Color segmentation using the C-Mean algorithm. This means segmenting the image into two classes; skin and non-skin regions, using neighborhood information to force the algorithm to create regions.

We addressed the problem of identifying skin-color and we adapt a spatial data mining method to this task and integrate with a segmentation method to identify significant skin-color regions in an image.

Our results suggest that skin color is a more powerful cue for detecting people in unconstrained imagery. Our solution showed its effectiveness, scoring a 96.1% classification accuracy rate on MYL test dataset and this method is used as a powerful tool for mining skin images [24,25].

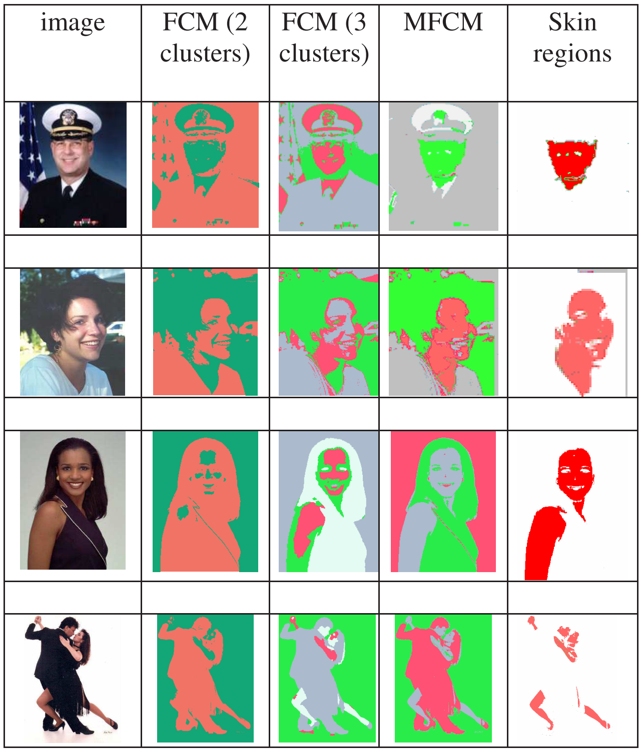

Figure 3 – Skin-segmented images showing the effectiveness of our approach on various ethnic races, lighting conditions and complex background

ACKNOWLEDGMENT

Special thanks goes to our students Mohamed Hammami and David Hodiesne who carried out the most of the experiments described in this paper. This work was partly funded and carried out within the French regional Comscope project.

REFERENCES

[1] M. Hammami, Y. Chahir, L. Chen, D. Zighed, Janvier 2003. "Détection des régions de couleur de peau dans l'image" revue RIA-ECA vol 17, Ed.Hermès, ISBN 2-7462-0631-5, pp.219-231.

[2] R. Deriche, Using Canny's criteria to derive a recursively implemented optimal edge detector, International Journal of Computer Vision, 1 (1987) 167-187

[3] S. Zhu , A. Yuille, Region competition : Unifying Snakes, Region Growing, and Byes/MDL for Multiband Image segmentation", IEEE Transactions on Pattern Analysis and Machine Intelligence, 18 (1996) 884-900.

[4] J. Morel, S. Soleimini, Variational Methods in Image Segmentation. Progress in Nonlinear Differential Equations Equations and their application, Birkhaüser 1995.

[5] M. Kass, A. Witkin, D. Terzopoulous, Snakes : Active Contour Models", International Journal of Computer Vision, 1, (1988) 321-331,.K. Sobottka and I. Pitas. Looking for faces and facial features in color images. Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications, Russian Academy of Sciences, 1996

[6] PEER, P., KOVAC, J., AND SOLINA, F. 2003. Human skin colour clustering for face detection. In *submitted to EUROCON 2003 – International Conference on Computer as a Tool*.

[7] AHLBERG, J. 1999. A system for face localization and facial feature extraction. Tech. Rep. LiTH-ISY-R-2172, Linkoping University.

[8] FLECK, M., FORSYTH, D. A., AND BREGLER, C. 1996. Finding naked people. In *Proc. of the ECCV*, vol. 2, 592–602.

[9] JORDAO, L., PERRONE, M., COSTEIRA, J., AND SANTOS-VICTOR, J., 1999. Active face and feature tracking. In *Proceedings of the 10th International Conference on Image Analysis and Processing*, 572–577.

[10] GOMEZ, G., AND MORALES, E. 2002. Automatic feature construction and a simple rule induction algorithm for skin detection. In *Proc. of the ICML Workshop on Machine Learning in Computer Vision*, 31–38.

[11] ZARIT, B. D., SUPER, B. J., AND QUEK, F. K. H. 1999. Comparison of five color models in skin pixel classification. In *ICCV'99 Int'l Workshop on recognition, analysis and tracking of faces and gestures in Real-Time systems*, 58–63.

[12] SCHUMEYER, R., AND BARNER, K. 1998. A color-based classifier for region identification in video. In *Visual Communications and Image Processing 1998, SPIE*, vol. 3309, 189–200.

[13] SORIANO, M., HUOVINEN, S., MARTINKAUPPI, B., AND LAAKSONEN, M. 2000. Skin detection in video under changing illumination conditions.In *Proc. 15th International Conference on Pattern Recognition*, vol. 1, 839–842.

[14] SIGAL, L., SCLAROFF, S., AND ATHITSOS, V. 2000. Estimation and prediction of evolving color distributions for skin segmentation under varying illumination. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 2, 152–159.

[15] CHAI, D., AND BOUZERDOUM, A. 2000. A bayesian approach to skin color classification in ycbcr color space. In *Proceedings IEEE Region Ten Conference (TENCON'2000)*, vol. 2, 421–424.

[16] JONES, M. J., AND REHG, J. M. 1999. Statistical color models with application to skin detection. In *Proc. of the CVPR '99*, vol. 1, 274–280.

[17] D.A.Zighed et R.Rakotomala, 1996. A method for non arborescent induction graphs. Technical report, Laboratory ERIC, University of Lyon2.

[18] M. Hammami, "Modèle de peau et application à la classification d'images et au filtrage des sites web", Thesis , Ecole Centrale de Lyon, 2005

[19] J.C. Bezdek, Pattern Recognition With Fuzzy Objective Function Algorithms. New York: Plenum Press, 1981.

[20] J. Bezdek, "A convergence theorem for the fuzzy ISODATA clustering algorithms," IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-2, pp. 1–8, 1980.

[21] H. Frigui, and R. Krishnapuram,, A Robust Competitive Clustering Algorithm With Applications in Computer Vision, IEEE Trans. on pattern analysis and machine intelligence, Vol. 21, N0. 5, 1999

[22] N. Boujemaa , "Generalized Competitive Clustering for Image Segmentation", 19th International Meeting of the North American Fuzzy Information Processing Society NAFIPS 2000 Atlanta, 2000

[23] J.-S. R. Jang, C.-T. Sun, E.Mizutani, Neuro-Fuzzy and Soft Computing, p (426-427)Prentice Hall, 1997

[24] M. Hammami, Y. Chahir and L. Chen , Web Filtering Engine Combining Textual, Structural and Visual Content-based Analysis , IEEE TKDE, 2005 (To appear) …

[25] M. Hammami, Y. Chahir, L. Chen : "WebGuard : Web based Adult Content Detection and Filtering System..", IEEE/WIC Int. conf. On Web Intelligence, pp.574-578. 2003.