BULETINUL INSTITUTULUI POLITEHNIC DIN IAŞI
Publicat de
Universitatea Tehnică „Gheorghe Asachi” din Iaşi
Tomul LVII (LXI), Fasc. 6, 2011
Secţia
ELECTROTEHNICĂ. ENERGETICĂ. ELECTRONICĂ

SKIN SEGMENTATION OF FACIAL IMAGES IN THE COMPRESSED DOMAIN WITH APPLICATIONS IN COSMETICS

BY

CAMELIA FLOREA[1,*], MIHAELA GORDAN[1], RADU ORGHIDAN[1], AUREL VLAICU[1], ANDREA BOTTINO[2] and ALDO LAURENTINI[2]

[1]Technical University of Cluj-Napoca
[2]Polytechnic University of Torino, Italy

Received, May 31, 2011
Accepted for publication: July 16, 2011

Abstract. In this paper we present a fast and accurate solution for skin segmentation of facial images in the JPEG compressed domain, in order to classify faces for further cosmetics applications. Within the increasing of image resolution in the past few years, the requirements for large storage space and fast processing (as developments directly in the compressed domain) became essential. The segmentation procedure used here is implemented on the compressed blocks, since this not only reduces the computation time by avoiding the decompression before processing, but also can benefit from direct measures of texture available *e.g*. in the Discrete Cosine Transform (DCT) coefficients domain and on the de-correlation of colour (YUV colour representation). This almost complete, yet reduced dimensionally, feature space makes easier the training and implementation of rather complex classifiers such as the Bayesian classifier with class probabilities modeled by Gaussian mixtures used here. Results on PUTFACE database have demonstrated the accuracy and robustness of the proposed approach.

Key words: compressed domain; image processing; Discrete Cosine Transform (DCT); skin segmentation, cosmetics.

[*] Corresponding author: *e-mail*: Camelia.Florea@com.utcluj.ro

1. Introduction

The cosmetics industry presents one of the highest areas of consumer spending. Looking lovely is not limited to wearing the right clothes, gadgets or accessories; it is also needed to wear the right make-up especially the right colours. The best way to know how to choose the right colours in your make-up is to use an application for face analysis (Bhatti *et al*., 2010; Rahman *et al*.,2010), to find the colours that fit your skin tone, eye colour and hair.

Despite the apparent simplicity of the skin detection problem, the research for more accurate solutions to this problem is still ongoing. The problem becomes even more difficult when the variability of the skin appearance increases – as it can often be the case in cosmetics applications where the type of facial skin and individual variations are higher.

The most popular approach formulates the task of skin segmentation as a pixel classification problem – that is, we are only interested to assign each pixel to the class skin or non-skin – and in most cases, the pixels colour is considered to be the most salient feature for this task. However to define the best description of the skin colour as opposite to any other colour is non-trivial. Therefore, almost all popular approaches are based on some learning methods to derive the skin colour models. Among these, a frequent (yet still improved (Zhanyu & Arne, 2010; Zafarifar & Martiniere, 2010)) approach is based on Bayesian classification.

There are several variations in the Bayesian skin detection algorithms, varying from the simple form of pixel classification based on the RGB (Red-Green-Blue) colour vectors, with non-parametric probability models or with Gaussian models (Phung *et al*., 2005), to more complex approaches, in decorrelated colour spaces, with other probability models (Zhanyu & Arne, 2010), or with colour and texture features used separately in the classification (Zafarifar & Martiniere, 2010) – all aiming to obtain higher accuracy of the skin detection.

In this paper we present a solution for skin segmentation of facial images in the JPEG (Joint Photographic Experts Group) compressed domain in order to classify faces for further make-up application. The use of DCT coefficients in segmentation and classification of texture has been explored lately by several researchers (Ghouzali *et al*., 2008). There are reported some skin detection approaches that work on DCT domain of JPEG image or MPEG (Moving Picture Experts Group) video and classify each block according to its colour and some also according to its texture properties (Lee & Hayes, 2004; Zheng & Gao, 2005; Abdallah *et al*., 2007; Jinchang *et al*., 2008; Mohamed *et al*., 2008; Zhao *et al*., 2010; Shiwei *et al*., 2009). To review a few of the DCT-based algorithms to skin segmentation, we should mention the paper published by Ghouzali *et al*. (2008), which extracts local DCT features for each pixel using small spatial neighborhoods, and further fits these coefficients to a generalized Gaussian distribution using the maximum-likelihood criterion applied to a set of training skin samples. Lee & Hayes (2004) have studied the skin and non-skin regions separation using the Bayesian decision rule with a Gaussian mixture model of skin colour, the features being the DC coefficients of *U* and *V* components. The approach studied by Mohamed *et al*. (2008) also relies on skin-based colour features extracted from DCT for face detection, and validates the skin face candidate by neural networks. Zhao *et al*. (2010) have extracted colour and texture features of the image blocks from the entropy decoded DCT coefficients firstly. Then, data mining method, *i.e*. decision tree, is applied to establish the skin colour model to describe the relationship between the features of image blocks and the skin detection results.

In this paper, the implementation of the facial skin segmentation directly in the compressed domain is investigated. Unlike other similar approaches, we aim to use solely and directly the information from the JPEG blocks, *i.e*. the zig-zag ordered quantized DCT coefficients in the YUV colour space, as feature vectors for the classification process. In this way it is possible to take advantage of direct measures of texture available, *e.g*., in the DCT coefficients domain and on the colours de-correlation (YUV colour representation). This reduced dimensionally feature space makes easier the training and implementation of rather complex classifiers, as the Bayesian classifier with class probabilities modeled by Gaussian mixtures used here, with practically no need for any feature extraction. In this way, the approach uses colour information, but also local texture information, considering image representation directly in the JPEG format.

2. Principles and Benefits of Compressed Domain Image Processing

Since the resolution of images increased in the past few years, the requirements for large storage space and fast processing became essential. The most used image storage format is JPEG format (over 95% of images on Internet are stored in this format), by reason of the advantages offered by this one: low storage capacity needed and better performances in information transmission.

The basic steps used to compress/decompress the JPEG images are the following. First, the image is divided into 8 × 8 blocks, and each 8 × 8 block is individually processed. A DCT is applied on each block and the resulting DCT coefficients are quantized. Many small coefficients, usually the high frequency ones, are quantized to zero. The next step is to zig-zag scan the DCT matrix; in this way, the AC coefficients are capable of indicate frequency components in an ascending order. The last steps in the coder’s algorithm are the RLE (Run Length Encoding) and entropy coding (Huffman coding). In the decoder, the compressed image is decoded, de-quantized and IDCT-transformed.

There are two ways to process the JPEG compressed images (Fig. 1): (i) *the compressed domain processing* – no decompression/compression, but the algorithm must be formulated in the DCT image representation space; (ii) *the pixel level processing* – process the image after decompression, direct manipulation of the pixels is adopted, than recompress the processed image.

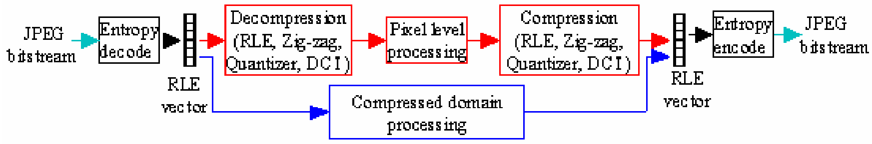

Fig. 1 – The two ways to process the image compressed JPEG.

If a segmentation algorithm can be formulated directly in the JPEG compressed domain, it is recommended to implement it, because a local texture measure is available directly, at no extra computational cost. The processing in compressed domain not only reduces the computation time by avoiding the decompression before processing, but also can benefit from JPEG reduced data space (the majority of the AC coefficients are zero after the quantization).

3. The Block Level Facial Segmentation Approach

Image segmentation represents the technique of partitioning an image into units which are homogeneous with respect to one or more characteristics. This can be done on JPEG pixels block level considering the local colour and texture information. There are many local features embedded in the DCT coefficients, reflecting colour and texture in spatial image windows: the DC coefficient represents the average colour in a block of pixels, while the AC coefficients reflect the variance of luminance and chrominance changes in pixels within the same block.

Particularly, in the case of skin detection, colour plays a significant role, since it provides information not available in the luminance component alone. The YUV colour space, used in the JPEG standard for colour image representation, is fortunately also a very good space for skin colour description (as mentioned in other similar applications (Zheng & Gao, 2005; Phung *et al*.,2002)). Therefore we express the segmentation process in this colour space, using the luminance and chrominance information from each image block. Furthermore, as the luminance, Y, and chrominance channels, U, V, are almost not correlated, we can use each of these features independently in the skin detection process, and combine the individual segmentation results in the end to obtain the overall skin segmentation. Thus, the proposed approach implies performing three block level segmentations in three features spaces: one formed by the zig-zag ordered quantized DCT coefficients of the Y component of each 8×8 pixels block, and the other two formed by the zig-zag ordered quantized DCT coefficients of the U and V components of each 8×8 pixels block – extracted directly from the JPEG image encoding. Let us further denote generically a zig-zag ordered vector of quantized DCT coefficients by $X[64\times1]$

(which is the maximum vector length), which can stand for Y_{dct} (in the case of the luminance component), U_{dct} or V_{dct} (in the case of the chrominance components).

As in general the skin colour distribution for a certain type of population tends to cluster in the YUV colour space, it is plausible to model it in each of the three feature spaces by a Gaussian model, as previously done by many researchers. Taking also into account that our segmentation performs on the block level (that is, in a vector, not scalar feature space), the suitable model for skin blocks distribution description is the Multivariate Gaussian model. For a given class, C_i , the Multivariate Gaussian model is completely described by the mean vector, m_i , and the covariance matrix, Σ_i. The probability density function of class C_i is computed by the Multivariate Gaussian Distribution

$$p(X \mid C_i) = \frac{1}{\sqrt{(2\pi)^{64} \left|\Sigma_i\right|}} \mathrm{e}^{-\frac{1}{2}(X-m_i)^T \Sigma_i^{-1} (X-m_i)} , \tag{1}$$

where $p(X \mid C_i)$ represents the probability of the vector $X[64 \times 1]$ to belong to class C_i, and the vector $X[64 \times 1]$ represents the zig-zag ordered quantized DCT coefficients of an 8×8 block.

The model parameters can be found from a set of representative data for the class, using the expressions for the mean vector, $m_i\,[64{\times}1]$, and covariance matrix, $\Sigma_i\,[64{\times}64]$

$$m_i = \frac{1}{N_i} \sum_{k=0}^{N_i - 1} X_{k,i} \ , \ \ \Sigma_i = \frac{1}{N_i} \sum_{k=0}^{N_i - 1} (X_{k,i} - m_i)(X_{k,i} - m_i)^T \ . \tag{2}$$

O b s e r v a t i o. If Σ_i is square and singular, then its inverse, Σ_i^{-1}, needed in (1), does not exists. This might be the case when Σ_i is sparse, as is the case when using zig-zag ordered quantized DCT coefficients. In these cases, the pseudo-inverse of Σ_i is used instead, as detailed in our previous work (Gordan *et al*., 2088).

The Multivariate Gaussian Model can further be employed in several probabilistic segmentation schemes; a common case is the Bayesian classifier, if the segmentation is defined as a binary classification of the data (in our case, of the pixels blocks) into skin or non-skin – the class skin being the above denoted class C_i, and non-skin – another complementary class, C_j. A simpler case is just the comparison of the skin probability given by (1) with some experimentally selected threshold.

In the skin segmentation approach, the above described feature vectors are extracted from the JPEG facial image bit stream after Huffman decoding. Each 8 × 8 pixels block is represented in the compressed domain actually by significantly less than 64 data, because most of the coefficients in the DCT domain are zero after the quantization step.

The algorithm is divided into two main phases: the training phase and the test phase. In *the training phase* of the algorithm the statistical properties of the classes used by the classifier are determined using ground-truth images. As a result the mean values and the covariance matrices (as well as their pseudo-inverses) are found for class skin in the *Y*, *U* and *V* feature spaces.

In *the test phase* of the algorithm the actual segmentation of an image is performed. Each and every 8×8 block from the compressed image is considered and the blocks are processed for classification. The classification is performed in the JPEG compressed domain, at block level, and all components (*Y*, *U* and *V*) are used to compute the probabilities of block belonging to skin region. Under the assumption that colour components and luminance information are not correlated, the joint class probabilities are the product of the luminance and colour probabilities, and the final decision rule is

$$\begin{aligned} &p(YUV_{\mathrm{dct}} \mid C_{\mathrm{skin}}) = p(Y_{\mathrm{dct}} \mid C_{\mathrm{skin}})\, p(U_{\mathrm{dct}} \mid C_{\mathrm{skin}})\, p(V_{\mathrm{dct}} \mid C_{\mathrm{skin}}), \\ &\begin{cases} p(YUV_{\mathrm{dct}} \mid C_{\mathrm{skin}}) \geq t, \ \text{skin region}, \\ p(YUV_{\mathrm{dct}} \mid C_{\mathrm{skin}}) < t, \ \text{non-skin region}, \end{cases} \end{aligned} \tag{3}$$

where: $p(YUV_{\mathrm{dct}} \mid C_{\mathrm{skin}})$ represents the probability of the 8×8 block to belong to class skin; Y_{dct}, U_{dct} and V_{dct} represent the zig-zag ordered quantized DCT coefficients of the 8×8 block from luminance component, *Y*, and, respectively, from chrominance components, *U* and *V*; *t* is the threshold value.

4. Implementation and Results

To verify the functionality of the proposed image segmentation algorithm on facial images, we implemented software the presented algorithm in the form of a Windows application. The program was tested on PUTFACE database which contain images of 100 persons in different positions stored as JPEG files.

There are two versions implemented for the segmentation process: one which leads to a soft segmentation, being based just on the skin probability estimation, and a second based on a hard thresholding applied on the skin probability map. For each facial image (Fig. 2 *a*) we obtained the skin probability map (Fig. 2 *b*) which contains the probability of each 8 × 8 block to belong to the skin region. The extracted skin regions are represented in Fig. 2 *c*.

Fig. 3 presents the classifications results using the segmentation in the compressed domain taking into account only the DC coefficients, (approach similar to the pixel level algorithm). Fig. 2 deems the whole DCT coefficients

(DC and AC coefficients) for segmentation. Figs. 2 and 3 show that highly accurate skin segmentation is performed using the whole DCT coefficients.

a

b

c

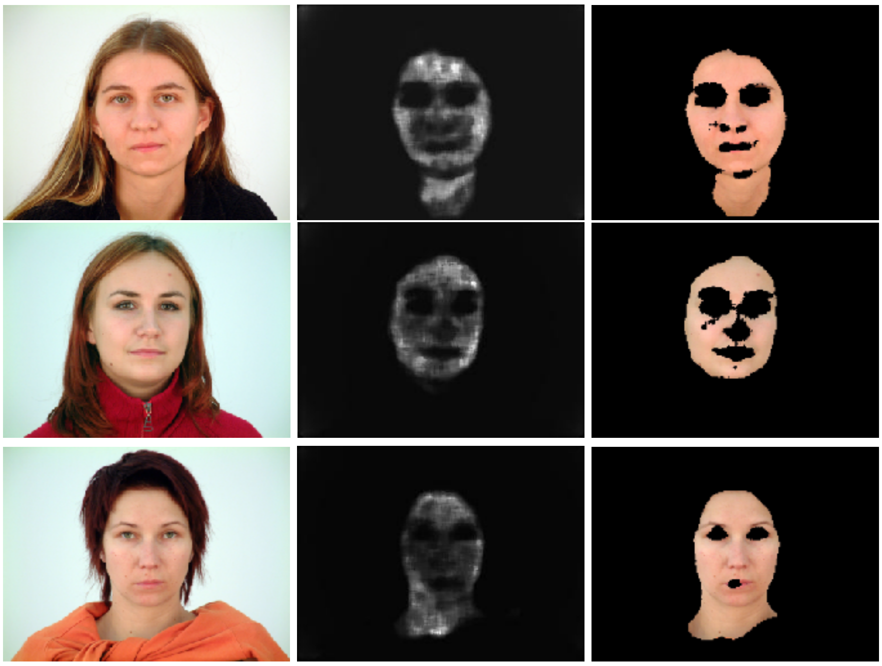

Fig. 2 – *a* – Original facial images; *b* – skin probability map; *c* – regions detected as skin.

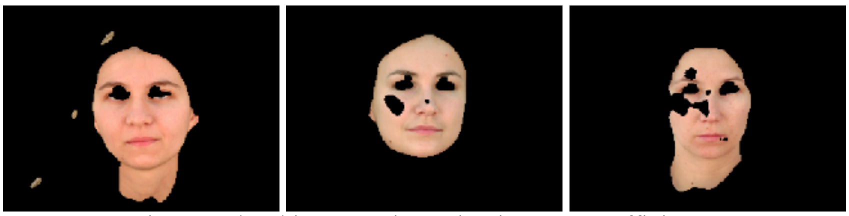

Fig. 3 – The skin extraction using just DC coefficients.

5. Conclusions

We propose and present a method for the segmentation of facial images, assumed to be represented in the JPEG compressed domain. Our goal was to provide an accurate tool for the identification of the skin regions, of reduced computational complexity – benefit that can be achieved when deriving and implementing segmentation algorithms directly on the compressed domain data

– in our case, the RLE vectors of quantized DCT coefficients in colour images. The approach was experimentally verified on PUTFACE database and proves good segmentation results.

Acknowledgment. This paper was supported by the project "Development and Support of Multidisciplinary Postdoctoral Programmes in Major Technical Areas of National Strategy of Research – Development – Innovation" 4D-POSTDOC, contract no. POSDRU/89/1.5/S/52603, project co-funded by the European Social Fund through Sectoral Operational Programme Human Resources Development 2007-2013.

REFERENCES

Abdallah S., Abdallah A., Lynn A., Mohamad A.E., *A New Face Detection Technique Using 2D DCT and Self Organizing Feature Map*. Proc. of World of Sci., Engng. a. Technol., **24**, 15-19 (2007).

Bhatti N., Baker H., Chao H., Clearwater S., Harville M., Jain J., Lyons N., Marguier J., Schettino J., Süsstrunk S., *Mobile Cosmetics Advisor: An Imaging Based Mobile Service***.** IS&T/SPIE Electronic Imaging: Multimedia on Mobile Devices, **7542**, 2010.

Ghouzali S., Hemami S., Rziza M., Aboutajdine D., Mouaddib E.M., *A Skin Detection Algorithm Based on Discrete Cosine Transform and Generalized Gaussian Density*. 15th Internat. Conf. on Image Process. (ICIP), San Diego, CA, Oct. 12-15, 2008, 605-608.

Gordan M., Popa C., Nagy G., Meza S., Vlaicu A., Mircea P., *Bayesian Segmentation of Hepatic Biopsy Colour Images in the JPEG Compressed Domain*. The 8th Internat. Conf. on Signal Process., Comput. Geom. a. Artif. Vision, Rhodes, Greece, Aug. 20-22, 2008, 140-145.

Jinchang R., Juan C., Jianmin J., Stan S.I., *Knowledge-Supported Segmentation and Semantic Contents Extraction from MPEG Videos for Highlight-Based Annotation, Indexing and Retrieval*. Adv. Intell. Comp. Theories a. Appl., **5226**, 2008, Springer, Berlin, 258-265.

Lee S., Hayes M., *A Simple and Fast Colour-Based Human Face Detection Scheme for Content-Based Indexing and Retrieval*. Internat. Conf. on Image Process. (ICIP), Singapore, Oct. 24-27, 2004, 701-704.

Mohamed A., Ying W., Jianmin J., Ipson S., *Face Detection Based Neural Networks Using Robust Skin Colour Segmentation*. Internat. Multi-Conf. on Syst., Signals a. Devices, Amman, Sept. 26, 2008, 1-5.

Phung S.L. Bouzerdoum A., Sr. Chai D., *Skin Segmentation Using Color Pixel Classification: Analysis and Comparison*. IEEE Trans. on Pattern Anal. a. Machine Intell., **27**, *1*, 148-154 (2005).

Phung S.L., Bouzerdoum A., Chai D. *A Novel Skin Color Model in YCbCr Color Space and Its Application to Human Face Detection*. Internat. Conf. on Image Process., **1**, Rochester, New York, USA, Sept. 22-25, 2002, 289-292.

Rahman A., Tran T., Hossain T., *Augmented Rendering of Makeup Features in a Smart Interactive Mirror System for Decision Support in Cosmetic Products Selection*. IEEE 14th Internat. Symp. on Distrib. Simul. A. Real Time Appl., Fairfax, VA, Oct. 17-20, 2010, 203-206.

Shiwei Z., Zhuo L., Xiao Z., Shen L., *A Data-Mining Based Skin Detection Method in JPEG Compressed Domain*. Sixth Internat. Conf. on Fuzzy Syst. a. Knowledge Discovery, Tianjin, Aug. 14-16, 2009, 297-301,.

Zafarifar B., Martiniere A., *Improved Skin Segmentation for TV Image Enhancement, Using Color and Texture Features*. Conf. on Cons. Electron., Digest of Techn. Papers Internat., Las Vegas, Nevada, Feb. 22, 2010, 373-374.

Zhanyu M., Arne L., *Human Skin Color Detection in RGB Space with Bayesian Estimation of Beta Mixture Models*. 18th Europ. Signal Process. Conf., Aalborg, Denmark, Aug. 23-27, 2010, 1204-1208.

Zhao S., Zhuo L., Wang S.L., Xiaoguang S.L., *Pornographic Image Recognition in Compressed Domain Based on Multi-Cost Sensitive Decision Tree*. 3rd IEEE Internat. Conf. on Comp. Sci. a. Inform. Technol., Australia, July 9-11, 2010, 225-229.

Zheng Q.F., Gao W., *Fast Adaptive Skin Detection in JPEG Images*. Adv. in Multimedia Inform. Process., Lecture Notes in Computer Sci., **3768**, 2005, 595-605.

DETECŢIA ZONELOR DE PIELE DIN IMAGINI FACIALE ÎN DOMENIUL COMPRIMAT PENTRU APLICAŢII COSMETICE

(Rezumat)

Se propune o soluţie performantă ca viteză şi acurateţe de detecţie a pielii din imagini faciale direct în domeniul comprimat (JPEG), necesară ca pas iniţial în clasificarea feţelor pentru aplicaţii cosmetice. Dezvoltarea unor metode rapide de prelucrare în domeniul comprimat este impusă de achiziţia şi stocarea imaginilor în format comprimat JPEG. Metoda de segmentare propusă operează direct asupra blocurilor comprimate JPEG, ceea ce asigură atât eficienţa sa numerică (prin evitarea decompresiei), cât şi o acurateţe sporită, deoarece coeficienţii transformatei cosinus dicrete furnizează atât informaţie de culoare, cât şi de textură în spaţiul de culoare YUV (care asigură decorelarea culorilor faţă de spaţiul culorilor primare). Acest spaţiu de trăsături furnizat implicit de codarea JPEG a blocurilor de imagine oferă o reprezentare compactă, cu o foarte bună aproximare, ceea ce facilitează implementarea unor clasificatoare instruibile performante, cum este şi clasificatorul Bayesian cu modelarea probabilităţilor claselor prin mixturi Gaussiene. Rezultatele experimentale obţinute pe baza de date PUTFACE demonstrează o rată de detecţie corectă a pielii prin metoda propusă, superioară altor implementări înrudite.