COLOUR SEGMENTATION

Petr Hujka, Petr Berka

Brno University of Technology, FEKT, Department of Telecommunications
Purkyňova 118, Brno 612 00, Czech Republic, {hujka,xrihak00, xuchyt00}@feec.vutbr.cz

Colour of pixel represents significant advantage in the field of segmentation of human skin. This segmentation can be used as input for many other applications, where it is requested to work only with regions containing skin samples. With an information about this regions, the whole image can be represented as binary image and processed with the dependence on the goal of other tasks.

Keywords: skin, segmentation, mahalanobis

1. INTRODUCTION

Colour of pixel represents significant advantage in the field of segmentation of human skin. If we consider a huge set of samples of pixels, that contain skin. This set can be obtained by manually selection, or with using some automatic technique. With this huge set of samples and computed threshold, we can say if inspected pixel is skin or not. Situation is not as easy as it looks like. Colour of human skin can change in dependence on the lighting, race of the specie, etc. Conditions of image capturing are always different and our task is to minimize the influence of this condition to segmentation process. The most disturbing conditions are the lighting conditions. In ideal case, we can separate them or suppress them. This task can be done by choosing suitable colour model.

2. COLOUR SPACES

Commonly used colour system, RGB colour space, which contains from three elements:

- R component (Red colour)
- G component (Green colour)
- B component (Blue colour)

This space is not applicable for segmenting, because information about luminance is spread in all three components and has significant influence to other components. Because we don't know the properties of light conditions, it's required to separate them. This can be done by using the Chrominance spaces. These spaces are using components, where the luminance is one of them or is significantly suppressed.

a) Normalized RGB Space

If we'll normalize components of RGB space, we will obtain rgb colour space, where the influence of luminance is suppressed. Normalization is done by using the following equations:

$$r = \frac{r}{R+G+B}, \quad (2.1)$$

$$g = \frac{g}{R+G+B}, \quad (2.2)$$

$$b = \frac{b}{R+G+B}. \quad (2.3)$$

It's evident from this equations, that the $r+g+b=1$. Normalized colours can be then represented only with two components, because for example r component can be obtained as:

$$r = 1 - g - b. \quad (2.4)$$

On the following image is illustrated 2 dimensional histogram in r-g subspace of set of human skin's samples rgb space.

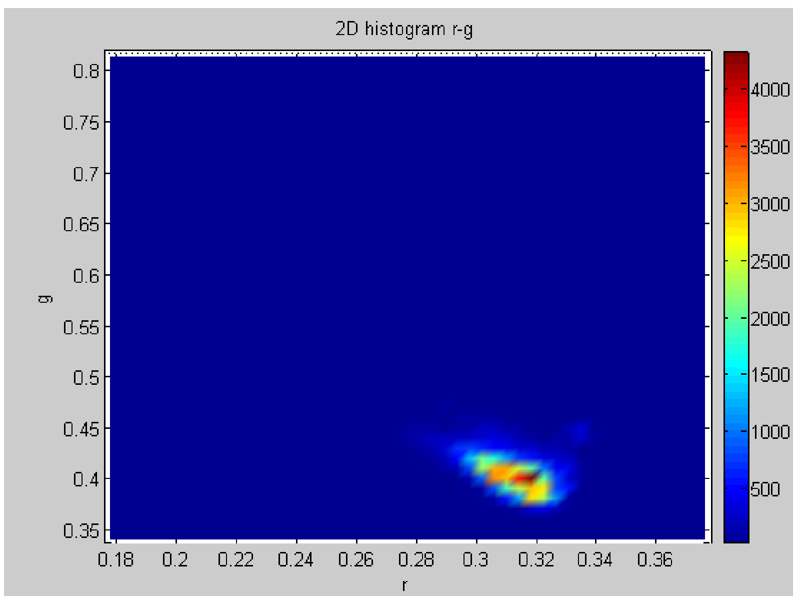

Fig. 1
2D histogram of samples of human skin's in r-g subspace

It's possible to see in Fig. 1 that all samples are in small cluster in around the mean value. This cluster has elliptical shape. If we analyze other two subspaces, we can find that they are very similar.

b) YCbCr Colour Space

Next inspected chrominance colour space is YCrCr space. This space is used in the field of digital video. Luminance component is there represented by one dimension of this system, by Y component. The information about colour is stored in the other two components, Cb and Cr. Cb component is value of difference between the reference value and blue component. Cr is then difference between reference value and red component.

Transformation from RGB colour space to YCbCr colour space is defined by:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65{,}481 & 128{,}553 & 24.966 \\ -37{,}797 & -74{,}203 & 112 \\ 112 & -93{,}786 & -18{,}214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (2.5)$$

Following image shows transformations of human skin's sample pixels to Cb-Cr subspace.

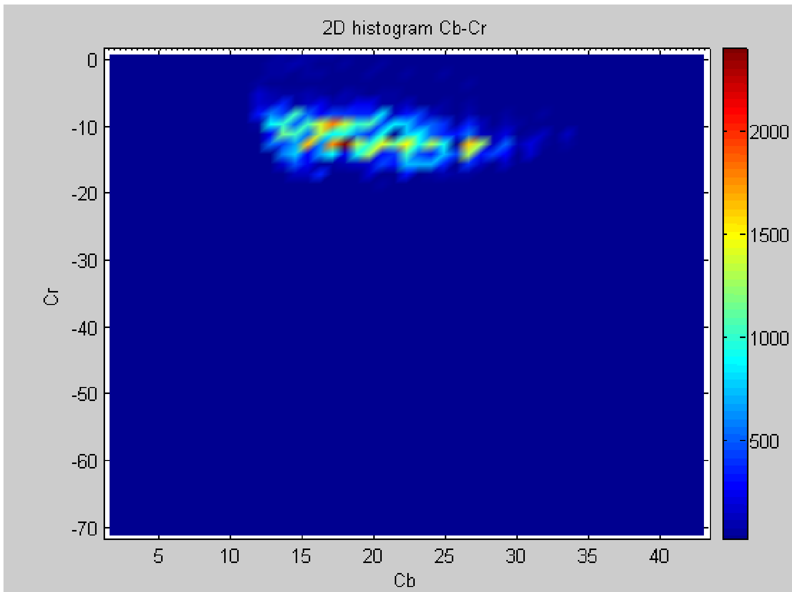

Fig. 2
2D histogram of samples of human skin's in Cb-Cr subspace

From figure 2, it's possible to see, that the samples are clustered again to small object of elliptical body.

c) <u>HSI Colour Space</u>

The most natural Colour space in the meaning of human perception is the HSI colour space. When human sees a colour object, he describes it as a combination of base colour (blue, red, yellow, etc…), saturation (the amount of white colour added to base colour) and intensity of base colour. Base colour is represented by angle in the range 0-360°. Colours are linear and form a closed circle.

Transformation from RGB space to HIS space is given by following equations:

$$H \begin{cases} \theta & ,if\ B \le G \\ 360 - \theta, & if\ B > G \end{cases}, \quad (2.6)$$

where

$$\theta = \cos^{-1}\left(\frac{\frac{1}{2}\left[(R-G)+(R-B)\right]}{\left[(R-G)^2+(R-B)(G-B)\right]^{\frac{1}{2}}}\right). \tag{2.7}$$

Saturation component is given by:

$$S = 1 - \frac{3}{(R+G+B)}\left[\min(R,G,B)\right]. \tag{2.8}$$

Finally, intensity:

$$I = \frac{1}{3}(R+G+B). \tag{2.9}$$

Following image shows transformations of human skin's sample pixels to H-S subspace.

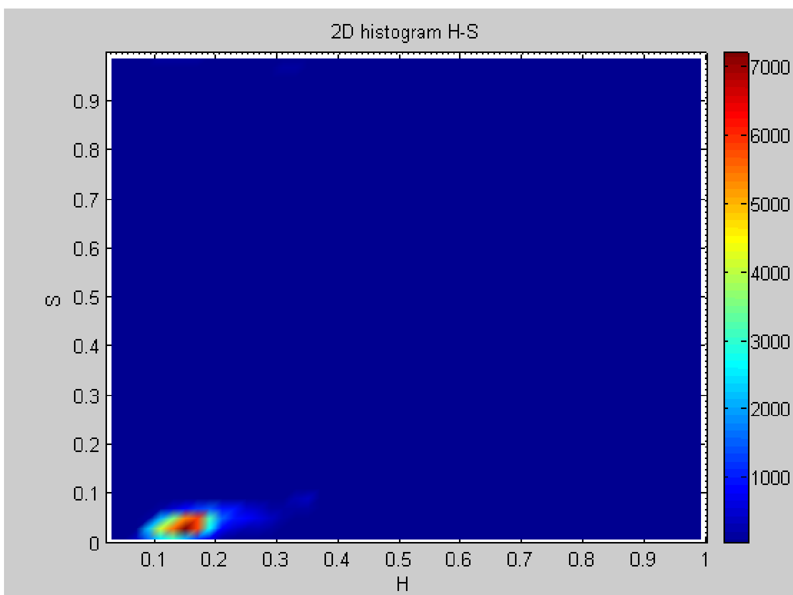

Fig. 3
2D histogram of samples of human skin's in H-S subspace

From figure 3, it's possible to see, that the samples are clustered again to small object of elliptical body.

3. STATISTICAL APPROACH

In previous chapter, we described the method of suppressing of different conditions. This chapter speaks about seeking of pixels and regions, what we are interested at.

In the following text, the samples of skin will be considered as multi-dimensional random value ξ, which is exactly defined by it's composite distribution function F($\boldsymbol{x}$), which is defined as probability, that components of vector ξ will be smaller then components of non-random vector $\boldsymbol{x}$:

$$F(\boldsymbol{x}) = P(\xi \le \boldsymbol{x}) . \qquad (3.1)$$

In the previous chapter, we found that the two dimensional distribution of samples of human skin's samples in chrominance spaces has elliptical body. It's composite density of probability can be then defined as:

$$f(\boldsymbol{x}) = (2\pi)^{-\frac{m}{2}} (\det \boldsymbol{C})^{-\frac{1}{2}} \exp\left(-\frac{1}{2} (\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{C}^{-1} (\boldsymbol{x}-\boldsymbol{\mu}) \right). \qquad (3.2)$$

Symbol det$\boldsymbol{C}$ describes determinant of covariance matrix $\boldsymbol{C}$ and symbol $\boldsymbol{x}^{\mathrm{T}}$ is transposition of vector $\boldsymbol{x}$. Parameters of this distribution are mean value vector μ and covariance matrix $\boldsymbol{C}$.

Vector of mean values μ, computed from multidimensional variety of size n, which is defined by n-tuple of m-dimensional vectors $\boldsymbol{x_i}$=($x_{i,1}$,..., $x_{i,m}$)$^{\mathrm{T}}$, which can be represented by matrix $\boldsymbol{X}$ of size n x m, is defined as:

$$\boldsymbol{\mu} = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{X}^T \boldsymbol{1}_n , \qquad (3.3)$$

where $\boldsymbol{1}_n$ is 1 x n vector, containing only ones.

Covariance matrix is defined as:

$$\boldsymbol{C} = \frac{1}{n-1} \sum_{i=1}^{n} (\boldsymbol{x}_i - \boldsymbol{\mu})(\boldsymbol{x}_i - \boldsymbol{\mu})^T . \qquad (3.4)$$

From (3.2) it's evident that we look for points, which doesn't yaw. Yawing data are data, which are far away from set of points with parameters μ and $\boldsymbol{C}$. This remoteness can be based on computation of Mahalanobis distance:

$$\boldsymbol{d} = \sqrt{(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{C}^{-1} (\boldsymbol{x}-\boldsymbol{\mu})} . \qquad (3.5)$$

If we suppose data with m-dimensional normal distribution, where μ and $\boldsymbol{C}$ were defined before, then yawing data are in region, where:

$$\boldsymbol{d} > \boldsymbol{d}_{krit} . \qquad (3.6)$$

The points, where $\boldsymbol{d}=\boldsymbol{d}_{krit}$ forms *m*-dimensional ellipsoid with the centre in μ. Area of yawing data, defined in (3.6) covers whole space E^m, except of this *m*-dimensional ellipsoid, which is surrounding vector of mean values.

Definition of $\boldsymbol{d}_{krit}$ can be done by computing of Mahalanobis distances for set of Human skin samples from vector of mean values of skin samples $\boldsymbol{d}_S$.

4. CONCLUSION

Colour of pixel represents significant advantage in the field of segmentation of human skin. This segmentation can be used as input for many other applications, where it is requested to work only with regions containing skin samples. With an information about this regions, the whole image can be represented as binary image and processed with the dependence on the goal of other tasks. We described effective method of human skin segmentation which is based on ststistical approach. This method suppresses the influence of disturbing condition of environment. The method was used to obtain region of interest of human skin for finding human faces in images.

5. REFERNCES

[1] HJELMAS, E., BOON, K. Face Detection: A Survey. In *Proceedings of the Computer Vision and Image Understanding.* 2001, vol. 83, p. 236-274. ISSN 1077-3142.

[2] BRUNELLI, R., POGGIO, T. Face recognition: Feature versus templates. In: *Proc. Of IEEE Trans. Pattern Anal. Mach. Intell*, 15, pages 1042–1052, 1993.

[3] TERILON, J.C., SHIHARAZI, M.S.: Comparative performance of Different Skin Chrominance Models and Chrominance Spaces for the Automatic Detection of Human Faces in Color Images. *Research report*, Kobe, Japan.