SCIENCE CHINA
Technological Sciences

November 2010 Vol.53 No.11: 2983–2988
doi: 10.1007/s11431-010-4132-z

A new method for face detection in colour images for emotional bio-robots

ZHANG Xu[1], ZHANG ShuJun[1,2*] & HAPESHI Kevin[1]

[1] *Computing Department, University of Gloucestershire, Park, Cheltenham, GL50 2RH, UK*
[2] *Key Laboratory of Terrain-Machine Bionics Engineering, Ministry of Education, Jilin University, Changchun 130025, China;*

Received May 18, 2010; accepted August 10, 2010

Emotional bio-robots have become a hot research topic in last two decades. Though there have been some progress in research, design and development of various emotional bio-robots, few of them can be used in practical applications. The study of emotional bio-robots demands multi-disciplinary co-operation. It involves computer science, artificial intelligence, 3D computation, engineering system modelling, analysis and simulation, bionics engineering, automatic control, image processing and pattern recognition etc. Among them, face detection belongs to image processing and pattern recognition. An emotional robot must have the ability to recognize various objects, particularly, it is very important for a bio-robot to be able to recognize human faces from an image. In this paper, a face detection method is proposed for identifying any human faces in colour images using human skin model and eye detection method. Firstly, this method can be used to detect skin regions from the input colour image after normalizing its luminance. Then, all face candidates are identified using an eye detection method. Comparing with existing algorithms, this method only relies on the colour and geometrical data of human face rather than using training datasets. From experimental results, it is shown that this method is effective and fast and it can be applied to the development of an emotional bio-robot with further improvements of its speed and accuracy.

face detection, skin colour model, eye detection, bio-robots

Citation: Zhang X, Zhang S J, Hapeshi K. A new method for face detection in colour images for emotional bio-robots. Sci China Tech Sci, 2010, 53: 2983–2988, doi: 10.1007/s11431-010-4132-z

1 Introduction

For the last decade, study of emotional and intelligent bio-robots has been a hot topic [1–7]. It is essential that such a robot has the intelligence of detecting human face. So, human face detection is an important research topic for image processing and intelligent bio-robots research and applications. It was initially considered as a key problem of the automatic face recognition system. The incipient face detection problem is based on the hypothesis that a frontal "mug shot" picture with strong constraints has been obtained [3]. In another word, it assumes there is only one face in the input image. However, in the real world pictures usually contain multiple human faces in a complex background. This raises new challenges to face detection technology.

Generally speaking, face detection is a pattern matching problem which includes two sub-problems: Firstly to determine if there are any faces within an input image; if the answer is yes, then to figure the face areas in the whole image [5]. Yang et al. [6] reviewed different face detection methods and classified them into two different catalogues: still image face detection and image sequence face detection. Actually, it is a unified problem if we notice that the video sequences are composed by a series of frames that can be considered as still images. Although some methods have

*Corresponding author (email: szhang@glos.ac.uk)

achieved good results, most face detection methods require datasets to train the classifiers [1, 7–9]. That means that it takes a relatively long time to detect faces if using a classifier. So this kind of methods cannot be effectively applied in some of on-line dynamic situations of practical applications such as bio-robots and on-line environment monitoring. Therefore, it is important to propose and develop a new method that does not need a classifier for human face detection in colour images.

2 The proposition of a new method for human face detection in colourimages

In this section, a new method will be proposed and developed for human face detection in colour images without relying on any training datasets. Figure 1 shows the diagram of the algorithm of the proposed method. Firstly, a luminance normalization step is used to adjust the luminance of the input image. Secondly, all skin colour regions are detected from the luminance-adjusted image using the skin colour model developed by Phuong-Trinh et al. [10]. These regions are considered as face candidates. Finally, each face candidate is verified by applying an eye detection method. The candidate would be considered as a face area if a pair of eyes is detected. Otherwise, it is not.

3 Skin segmentation

Colour is one of the distinctive properties of human skin. The advantage to detect skin by its colour is obvious because skin colour keeps as invariant even when rotation and occlusion occur. However, the colour is sensitive to lighting conditions, which means the luminance has significant effect on the pixels' values of colour images. Moreover, the statistic of skin colours is varied among different human races. Therefore, it is still a challenging task to research.

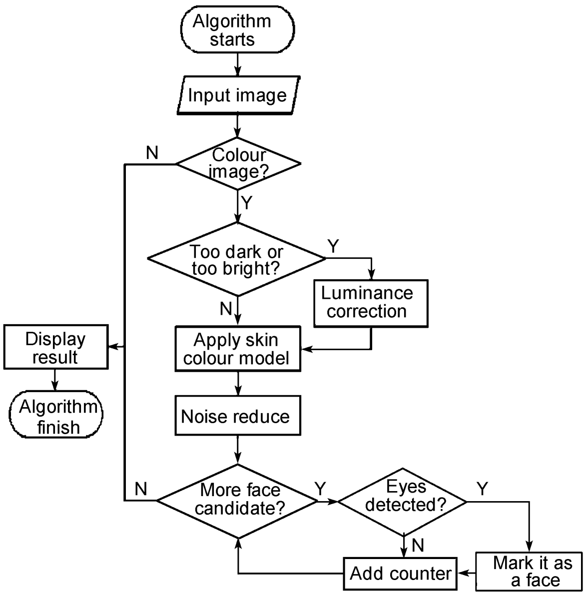

Figure 1 Face detection algorithm.

3.1 Colour spaces

Vezhnevets et al. [11] has compared several skin colour detection techniques which were carried out in different colour spaces. Researchers have proposed diverse detection methods in different colour spaces, such as RGB, sRGB, CMYK, HSV, YCbCr, YUV, YES, TSL, HSL, HIS, YIQ, and CIE [12]. It's hard to determine which colour space is the best choice for face detection. Although RGB colour space does not separate the luminance from chrominance, it is still the straightforward space to understand while this space, consisting of three components, R (red), G (green) and B (blue), is the most frequently used for displaying and storing colour image in industry. YCbCr is another popular colour space for face detection [13, 14]. In this paper, RGB space is used as the primary colour space in which the skin colour model is applied. Meanwhile, YCbCr colour space is used for correcting the luminance of the image.

3.2 Luminance normalization

Because the luminance of image will seriously affect the skin segmentation results, a luminance normalization step is necessary before applying skin colour model to make the algorithm robust. A typical method to deal with the luminance problem is called "reference white", which was proposed by Hsu et al. [14]. According to the method, the top 5% of luminance values in the image are considered as "reference white" if the number of these pixels is sufficient large (greater than 100). Then the R, G, and B components of the image are adjusted by three coefficients, so that the gray values of reference-white pixels are linear scaled to 255 (real white).

Although the reference white method is said to be useful for face detection in YCbCr colour space, it is not effective as expected to detect human face in the RGB colour space. Thus, a new luminance normalization method has to be developed to correct the luminance while the input image is too dark or too bright. Firstly, the image is converted from the RGB colour space to the YCbCr space according to the following equation [4]:

$$\begin{pmatrix} Y \\ Cb \\ Cr \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.0169 & -0.332 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{pmatrix} \times \begin{pmatrix} R \\ G \\ B \end{pmatrix}. \qquad (1)$$

In the YCbCr space, Y component, the luminance of the original image, is a gray-scale copy of the original image, and Cb and Cr components represent the chrominance of the image. Then the luminance normalization rule is defined as

$$\mathrm{Y}(x,y)=\begin{cases}\mathrm{Y}(x,y) & \text{if mean (Y)>120 and } \mathrm{R}\left(\dfrac{\mathrm{Y}}{\mathrm{CbCr}}\right)\in[0.9,1],\\ \left(\dfrac{\mathrm{Y}(x,y)}{\mathrm{R}\left(\dfrac{\mathrm{Y}}{\mathrm{CbCr}}\right)}\right), & \text{otherwise,}\end{cases} \tag{2}$$

where

$$\mathrm{R}\left(\frac{\mathrm{Y}}{\mathrm{CbCr}}\right)=\frac{\text{mean (Y)}}{2[\text{mean (Cb)}+\text{mean (Cr)}]}$$
$$=\frac{\dfrac{1}{m\times n}\sum_{i=0}^{m}\sum_{j=0}^{n}\mathrm{Y}(i,j)}{\dfrac{1}{2}\left(\dfrac{1}{m\times n}\sum_{i=0}^{m}\sum_{j=0}^{n}\mathrm{Cb}(i,j)+\dfrac{1}{m\times n}\sum_{i=0}^{m}\sum_{j=0}^{n}\mathrm{Cr}(i,j)\right)},$$

is the luminance ratio; m and n are the height and width of the image, respectively; Y, Cb and Cr are the three components in the YCbCr colour space. Because the picture may be taken at night when the background luminance value is small, the luminance needs to be normalized only if the mean of luminance is bigger than 120. In this satiation, the image is considered to be too dark if the luminance is smaller than 0.9; on the other hand, it is considered to be too bright if the luminance is greater than 1. The correction range [0.9, 1] is based on the experimental experiences.

3.3 Skin colour model

In this method, a skin colour model will be proposed based on the model developed by Phuong-Trinh et al. [10]. By comparing with two other skin models [13, 15], the model developed by Phuong-Tring defines the boundaries of skin colours explicitly in the RGB space. It is efficient to avoid retaining non-skin colours, for example yellow, white, orange, pink, red, wood-brown and sand-yellow. The decision rules of this skin model are defined as

$$\delta(P(x,y))=\begin{cases}1, & \text{if } \mathrm{S}(P(x,y)) \text{ is satisfied,}\\ 0 & \text{otherwise,}\end{cases} \tag{3}$$

where (x, y) are the coordinates of pixels in the image; $P(x, y)$ is a pixel value of colour; image condition $S(P(x, y))$ is listed in Table 1.

From Table 1, it can be seen that the skin model adjusts reasonable differences between R, G, and B components. If this model is only used to identify face areas, some non-face areas may also be falsely treated as a face area. The authors of this paper have done a numerous trials and tests of examining the values of skin areas selected manually using a mouse, and it has been found that most human being face's colours follow the rule of

$$\mathrm{R}>1.1\times\mathrm{G}>1.1\times\mathrm{B}, \tag{4}$$

which means R component is always the strongest one. So the skin model expressed in eq. (3) can be modified to a new skin model by applying the rule of eq. (4). This new skin model will be used in this paper.

After skin segmentation, connected skin areas should be labeled. Then the next step is "noise reduction" to erase out the areas which are smaller than a given threshold as defined in eq. (5). To deal with images of different sizes, a self-adapting threshold is used according to the image size. This step has been proved to be more effective in processing by reducing unreasonable face regions.

$$\delta(S)=\begin{cases}1, & S\geqslant t,\\ 0, & \text{otherwise,}\end{cases} \tag{5}$$

where S is the area of detected regions in binary image and t is the self-adapting threshold computed with the size of the image.

3.4 Geometric restrictions

The skin colour model used in Section 3.2 is effective to detect skin region from the image. However, it's not easy to distinguish faces and other body parts such as arms and hands. Therefore, some geometric restrictions have to be applied to those regions to eliminate non-face candidates as soon as all skin regions are detected from the images.

Table 1 A set of conditions defining skin pixels. These conditions should be satisfied simultaneously

R	(R-G)	(G-B)	B	G
[70, 85]	[30, 55]	[−5, 35]	[20, 255]	[30, 255]
[86, 100]	[30, 60]	[−5, 40]	[30, 255]	[40, 255]

R	(R-G)	(G-B)	(R-B)	(R+B-2G)	G
[101, 150]	[0, 30]	[−10, 45]	[15, 75]	[−15, 285]	-
	[31,75]	[−5, 90]	[−255, 120]	[−20, 285]	[50, 255]

R	(R-G)	(G-B)	(R-B)	(R+B-2G)	B
[151, 200]	[15, 20]	[−5, 40]	[20, 255]	[−20, 285]	-
	[31, 85]	[−15,70]	[20, 255]	[0, 285]	[40, 255]

R	(R-G)	(G-B)	(R+B-2G)
[201, 255]	[5, 25]	[40, 70]	[−30, 285]
	[26, 100]	[0, 70]	[−15, 285]

Although Séguier [16] tried to locate elliptic skin region using Hough transform and considered them as faces, his method was less efficient to deal with the condition in which skin regions connected with each other. The model we proposed in Section 3.2 has a similar problem. To overcome this problem, another step has to be introduced to separate connected skin regions. It is normally that most previous studies use eq. (6) to calculate projection sums for each horizontal rows and vertical columns of skin pixel intensity map. By comparing these sums with width and height of the current skin area, the sub-areas with values of sums smaller than a given threshold are considered as concave regions. These regions will be removed; hence, the connected areas can be separated.

$$\begin{cases} H(i) = \sum_{j=0,i}^{n} B(i,j), \\ V(i) = \sum_{i=0,j}^{m} B(i,j), \end{cases} \quad i \in [0,m],\ j \in [0,n], \tag{6}$$

where $B(i, j)$ is the binary image of the current skin area; m and n are the height and width of the current skin area, respectively; $H(i)$ and $V(j)$ are the projection sum of horizontal rows and vertical columns, respectively.

After the connected regions are separated, each region is labeled and considered as an independent face candidate. Then, a series of morphology operators (i.e., erode, dilate, open, and close operators) are applied to the binary image to fill the holes in the detected areas. After investigating the properties of human faces, we found that the height-width ratios are never bigger than 2.5. So all the non-face skin regions should be eliminated based on several geometric conditions as follows

$$\delta(S) = \begin{cases} 1, & \dfrac{\max(H,W)}{\min(H,W)} < 2.5, \\ 0, & \text{otherwise}, \end{cases} \tag{7}$$

where S is the number of skin pixels belonging to the skin region rectangle, and H and W are the height and width of the skin region rectangle.

After a serious of tests using the method described above, all face areas can be identified. But some body parts with geometrical shape similar to the face cannot be removed. It has been particularly found that human hand is a difficult candidate to eliminate if only applying the method. Section 4 will describe a new method for eliminating the false face candidate.

4 Eye detection method

4.1 The algorithm

To eliminate false face candidates, an eye detection method is used. The flowchart of the eye detection algorithm used in this paper is illustrated in Figure 2.

4.2 Eye detection rules

Rule 1: The vertical location of eyes is near the peak point of the vertical projection sum of the difference image.

Because the pixels around eyes area are usually with lower values in the gray-scale image, the vertical position of eyes area can be located by computing the horizontal projection sum according to

$$\begin{cases} H(i) = \sum_{i,j=0}^{n} V(i,j), \\ V(i,j) = F(i,j) - I(i,j), \end{cases} \tag{8}$$

where $V(i, j)$ is the difference image; $I(i, j)$ is the gray-scale image of face candidate; $F(i, j)$ is the result obtained by applying close operation with morphological operator, which is a disk-shaped structure element with radius 5.

As shown in Figure 2, the eye area is always the highest peak of smoothed horizontal projection sum (green point in Figure 3(a)).

Rule 2: The eyes appear in the upper half in a face and the areas of eye blocks are within a certain range.

In any faces which are not rotated significantly, it is obvious that such a rule should be obeyed. By applying a threshold to the difference image, all potential eye areas remain as the white blocks in the binary image (Figure 4(d)). To make this method more robust, the binary image is processed by a morphological close operation with disk-shaped structure element with radius 3 (Figure 4(e)), and then blocks that are with small areas or in the lower half of the binary image will be eliminated. Finally, the remaining blocks would be considered as eyes pair (Figure 4(f)).

Rule 3: The eye blocks are of approximately round shape.

The height-width ratios of all remaining blocks will be computed. Then those blocks whose height-width ratios are greater than a given threshold will be eliminated. Furthermore, the possible eyes pair should have similar area.

After applying Rules 1–3, the algorithm will return false if there is no block left. Thus, the current face candidate will be considered as a “false face”. Otherwise, the algorithm continues by checking all possible eyes pairs to compare their similarity.

Rule 4: The distance between two eyes is about 1/4–1/2 of the width of face.

This rule is used for eliminating impossible eyes pair, which means only blocks with certain distance will be checked. This rule could reduce the processing time.

Next, each pair of remaining blocks is treated as a possible eyes pair. According to their centroids, we crop two circle areas with radius 20 in the gray-scale image of current face candidate. Then, their similarity is calculated according to eq. (9). If the similarity is greater than 0.5, their indices are saved in an array. After the loop, the pair with the highest similarity is considered as "real eyes" if the indices array is non-empty, and then the current face candidate is marked as a "real face". Otherwise, it is a "false face".

$$S=\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(A_{i,j}-\bar{A})(B_{i,j}-\bar{B})}{\sqrt{\left(\sum_{i=1}^{m}\sum_{j=1}^{n}(A_{i,j}-\bar{A})\right)^2\left(\sum_{i=1}^{m}\sum_{j=1}^{n}(B_{i,j}-\bar{B})\right)^2}},\quad(9)$$

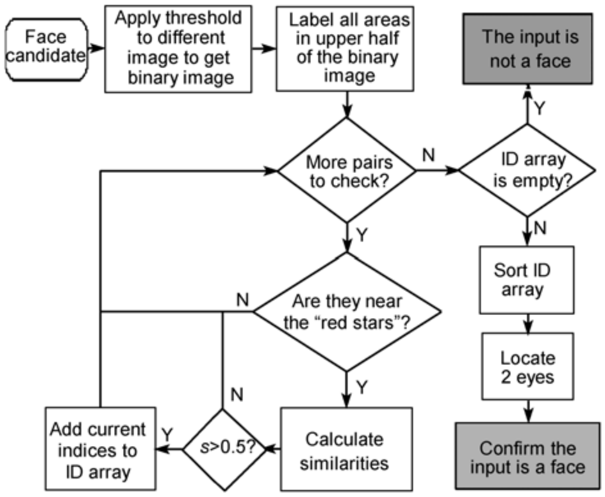

Figure 2 Eye detection method.

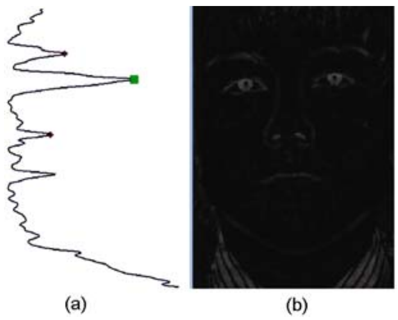

Figure 3 Horizontal projection of valley image Horizontal Projection (a); difference image (b).

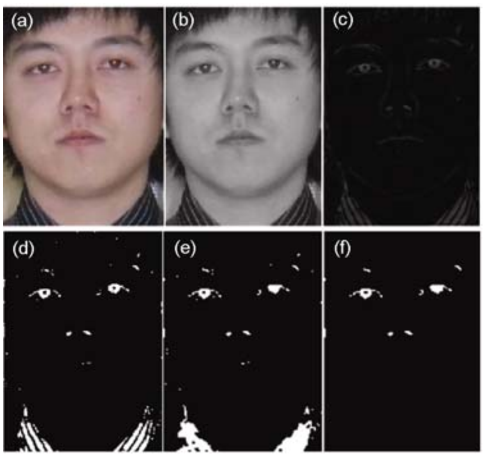

Figure 4 (a) Face candidate; (b) gray-scale image; (c) difference image; (d) binary image; (e) morphological close image; (f) noise reduced image.

where A and B are the two areas for comparison; $\bar{A}$ and $\bar{B}$ are the means of their pixels' values.

5 Experiment

Figure 5 shows a complete face detection example and the intermediate steps during the algorithm running.

The face detection method proposed in this paper was tested on 40 different images taken in various conditions. All images were with complex background and different lighting conditions. There were 65 faces included in the testing images. As shown in Table 2, our method detected 56 faces successfully, while 9 faces were missed and 4 false faces were detected.

As shown in Figure 6, the processing time for skin colour segmentation was less than 0.2 s for 500×400 pixel images. The total execution time for the same size is about 1 s.

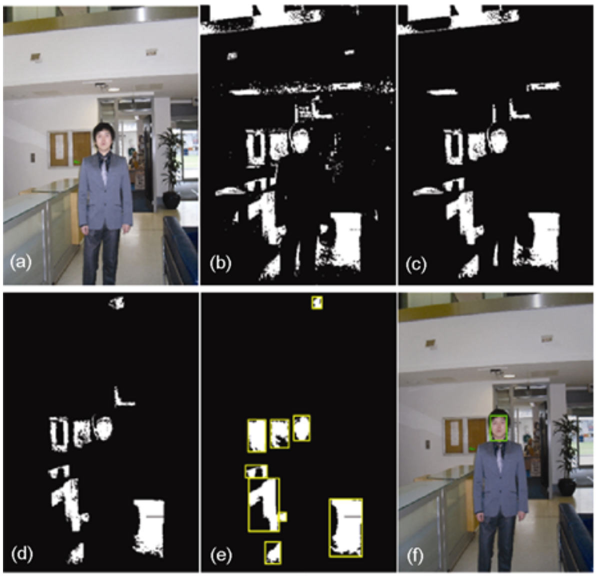

Figure 5 Experimental result. (a) Original image; (b) skin area detected; (c) noise reduced image; (d) geometric restriction; (e) face candidates; (f) detected face.

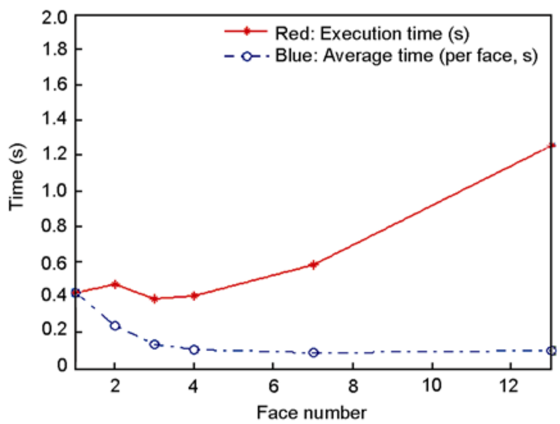

Figure 6 Execution time.

Table 2 Experiment results

	Images	Faces	Detected	Missed	False
Number	40	65	56	9	4
Rate			86.15%	13.85%	6.15%

However, the execution time for a 2800×3700 pixel image will increase to more than 14 seconds, because for a big image more possible eye blocks need to be checked for the similarity for each face candidate.

6 Conclusion

In this paper, a new face detection method has been proposed for detecting faces in colour images. It is a coarse-to-fine process: First locating all skin areas by applying skin colour model, and then examine each possible face candidate by eye detection algorithm. To search facial features within face candidate area rather than the whole image can reduce the computation complexity. By comparing to existing algorithms for face detection, the most important advantage of our method is no training data needed. The experiment results have shown that the method is robust and effective. Further work is necessary to optimize the eye detection rules to make this method fast enough to deal with big images.

1 Romdhani S, Torr P, Schölkopf B, et al. Computationally efficient face detection. In: the 8th International Conference on Computer Vision, Vancouver, BC, Canada, 2001. 697–700
2 Zhang S J, Hapeshi K, Bhattacharya A K. 3D modelling of biological systems for biomimetics. J Bionic Eng, 2004, 1: 20–40
3 Ballot J S S. Face Recognition Using Hidden Markov Models. Dissertation of Masteral Degree. Matieland: University of Stellenbosch, 2005
4 Berbar M A, Kelash H M, Kandeel A A. Faces and facial features detection in color images. In: Proceedings of Geometric Modeling and Imaging: New Trends (GMAI'06), London, England, 2006. 209–214
5 Hjelmas E, Low B K. Face detection: A survey. Comput Vision Image Understanding, 2001, 83: 236–274
6 Yang M H, Kriegman D, Ahuja N. Detecting faces in images: A survey. IEEE Trans Pattern Analysis Mach Intell, 2002, 24: 34–58
7 Terrillon J C, David M, Akamatsu H. Automatic detection of human faces in natural scene images by use of a skin color model and of invariant moments. In: the Third International Conference on Automatic Face and Gesture Recognition, Nara, Japan, 1998. 112–117
8 Rowley H A, Baluja S, Kanade T. Neural network-based face detection. IEEE Trans Pattern Analysis Mach Intell, 1998, 20: 23–38
9 Wu B, Ai H, Huang C, et al. Fast rotation invariant multi-view face detection based on real AdaBoost. In: the Sixth IEEE International Conference on Automatic Face and Gesture Recognition, Seoul, Korea, 2004. 79–84
10 Pham-Ngoc P T, Huynh Q L. Robust face detection under challenges of rotation, pose and occlusion [OL]. The Second Young Vietnamese Scientists Meeting 2007, [2 Jan, 2010]. http://www.fas.hcmut.edu.vn/webhn10/Baocao/PDF/YVSM_PhuongTrinh.pdf
11 Vezhnevets V, Sazonov V, Andreeva A. A survey on pixel-based skin color detection techniques. In: GraphiCon 2003, Moscow State University, 2003. 85–92
12 wikipedia.org. List of color spaces and their uses. Retrieved on: 10 Dec, 2009; from: http://en.wikipedia.org/wiki/List_of_color_spaces_and_their_uses.
13 Kovac J, Peer P, Solina F. Human skin color clustering for face detection [OL]. EUROCON 2003.The IEEE Region 8, 2003, 2: 144–148 [10 Nov, 2009]. http://lrv.fri.uni-lj.si/~peterp/publications/eurocon03.pdf
14 Hsu R L, Abdel-Mottaleb M, Jain A K. Face detection in color images. IEEE Trans Pattern Analysis Mach Intell, 2002, 24: 696–706
15 Lin C, Fan K C. A color-triangle-based approach to the detection of human face. In: Biologically Motivated Computer Vision, First IEEE International Workshop, BMVC 2000, Seoul, Korea, 2000. 195–200
16 Séguier R. A vary fast adaptive face detection system. In: International Conference on Visualization, Imaging and Image Processing, Marbella, Spain, 2004