Correction of high lighting using histogram matching for skin segmentation on white background images

Alex T. S. Carneiro  
alextcarneiro@gmail.com

Paulo C. Cortez  
cortez@lesc.ufc.br

Rodrigo C. S. Costa  
rodcosta@lesc.ufc.br

Laboratório de Engenharia de Sistemas de Computação - LESC  
Universidade Federal do Ceará - UFC  
Campus do PICI S/N, Bloco 723, Cx.Postal 6015, CEP 60455-970, Fortaleza - CE - Brazil

ABSTRACT
In this work, the histogram matching is used as a tool to improve the skin segmentation in images with white background and high illumination. Traditional skin detection algorithms realize the segmentation without doing any enhancement in the images and the proposed enhancement method solves this point. The results of this work demonstrate that the proposed method improves the quality of image segmentation at least 75%, when compared with manual segmentation.

Keywords
Skin segmentation, Image enhancement, Histogram matching

1. INTRODUCTION
Human Computer Interfaces (HCI) are becoming smarter to each year [26]. Some techniques used in HCI are intelligent sensors [25, 4, 24] and Computer Vision (CV) [22, 26, 8, 5, 23]. Many state of art systems based on CV need an additional device to make the object detection easier, but those systems are uncomfortable for users [5, 23]. So, systems that interact directly with human body are more indicated for common users.

Skin segmentation is a fundamental step in many Human Computer Interfaces that interact directly with human body, like gesture based. A key problem of this kind of algorithm is its dependence of illumination conditions [20, 27]. This work proposes a technique to correct the high illumination and to improve the skin segmentation in a white background, while the state of art doesn’t address this interesting aspect.

The proposed method consists in apply the histogram matching to change the original histogram to a controlled histogram [10]. This technique requires the knowledge of an established background. During the study, images with white background were used to specify the histogram. This work is part of a Brazilian Sign Language Recognition System.

This work is organized as follow. Section 3 presents and describes the background algorithm used in the work. The experimental setup is described in Section 4. The proposed method is explained in the Section 5. The results are presented in Section 6 and Section 7 presents the conclusion and comments.

2. RELATED WORKS
Skin segmentation is an important component in many systems for human-computer interaction or any system that should detect people, like systems for security or monitoring [2, 27, 21, 1, 12, 11, 3]. The more used approaches involve models of statistical analysis of the images [21, 1] or changes in the imaging devices [12, 3].

3. BACKGROUND
This work proposes a technique to improve skin segmentation and this section presents the used algorithms and describes how they work.

3.1 Histogram matching
The histogram of a digital image with gray levels in the range \[0, L - 1\] is a discrete function \(h(r_k) = n_k\), where \(r_k\) is the \(k\)-th gray level and \(n_k\) is the number of pixels in the image having the gray level \(r_k\) [10]

\[n_k = \frac{p(r_k)}{N},\]

for an image with \(N\) pixels and \(p(x)\) is the incidence’s probability of \(x\) in the image.

The histogram matching, or histogram specification, is an algorithm for change the histogram to any other shape using two functions, \(T(r)\) and \(G(z)\), as follow [10].

Given \(p_r(r)\) the original histogram, \(T(r)\) is given by

\[T(r) = \int_0^r p_r(w)dw,\]

and, given a desired histogram \(p_z(z)\), \(G(z)\) is given by

\[G(z) = \int_0^z p_z(t)dt,\]

where \(w\) and \(t\) are dummy variables.
Doing $G(z) = T(r)$, so the obtained transformation is

$$z = G^{-1}[T(r)].$$  \(4\)

So, for each pixel’s value \(r\) the histogram matching applies the transformation $z = G^{-1}[T(r)]$, and the result is an image which its histogram is similar to \(p_z(z)\) desired.

### 3.2 Color images models

In this work, the images are represented in color models because skin is better represented in color images [9]. Color images have three channels of color, usually RGB (Red - Green - Blue) the basic color components [10]. Other representations of colors are obtained using transformations over the image, the most commons are: YCbCr (Luminance - Chrominance blue - Chrominance red), CMY (Cyan - Magenta - Yellow) and HSV (Hue - Saturation - Value).

Each color representation enhances different features from the image, e.g. the YCbCr decomposes the image in one channel for bright (luminance) and two channels for colors (chrominance blue and chrominance red) [14, 7].

The RGB to YCbCr transformation is given by \([7]\)

$$Y = 0.299R + 0.587G + 0.114B,$$
$$Cb = 0.564(B - Y),$$
$$Cr = 0.713(R - Y).$$  \(5\)

The skin segmentation algorithm used in this work uses the components \(Cb\) and \(Cr\). Those components have no bright information, but they undergo influence of the light and it may do the segmentation worse. The most ones works in the literature use the components \(Cb\) and \(Cr\) in skin segmentation algorithms, presented in [9], [6] and [13].

### 3.3 Skin segmentation

Segmentation is a process of labeling in the image to identify desired regions on the image [10, 18]. In this work, the region desired is the human skin, in special the hand.

The algorithm used for skin segmentation is implemented like following. The color coding used is the YCbCr. The Cb and Cr components of each pixel are analyzed using the following Equation \([17]\)

$$77 \leq Cb \leq 127 \text{ and } 133 \leq Cr \leq 173$$  \(6\)

the Equation 6 characterizes an image thresholding in the channels Cb and Cr of the image.

This is a key algorithm for skin segmentation because its simplicity and different skin tones are grouped in a specified region, the difference between two, or more, persons is the absorption of light by the skin. The Figure 1 shows an image and the result of thresholding algorithm. However, the high lighting intensity may change the acqirement colors.

Observing the Figure 1, the algorithm segments dark and clear skins correctly, but it confuses skin with red regions (non-skin). Therefore, the test’s images were obtained in a controlled environment without red regions.

![Figure 1: Images for validation of the skin segmentation algorithm (a) original image and (b) image segmented.](image)

### 4. EXPERIMENTAL SETUP

This work was developed in a Notebook with the following configuration: Intel Core 2 Duo® processor, 2 GB of RAM memory. The images were captured using a VGA webcam (640 x 480 pixels) with focus control.

The softwares used in this work, for development of the algorithms and simulations, were the Scilab and the toolbox for image manipulating Scilab Image and Video Processing toolbox (SIVp), over the GNU/Linux Ubuntu 8.04 operating system. About the SIVp toolbox, it provides a set of functions of image manipulation and processing, for this work were used functions for read/write files and convert the color model RGB to YCbCr.

The SIVp toolbox adds a constant term of 128 in the Cb and Cr components, so the used transformation from RGB to YCbCr is

$$Y = 0.299R + 0.587G + 0.114B,$$
$$Cb = 0.564(B - Y) + 128,$$
$$Cr = 0.713(R - Y) + 128.$$  \(7\)

The softwares were chosen because they are free software and compatible with the proposed application. All softwares used in this work are available for freely download in: http://www.scilab.org/ and http://www.ubuntu.com/

### 5. PROPOSED METHOD

The proposed method consists in specify the desired histogram $p_z(z)$ presented in the Equation 3.

#### 5.1 Histogram of reference

The images have two regions: hand and background. So, the histogram presents two concentrations of pixels: one to the hand and other to the background, like showed at the Figure 2. The hand’s concentration is called $G_h(Cb, Cr)$ and the background’s concentration is called $G_b(Cb, Cr)$.

Analyzing the histograms of the Figures 2(c) and 2(d), the conclusion is that the greater concentration (background) is about ten times higher than the other.

The result of skin segmentation algorithm applied to image of the Figure 2(a) is shown in Figure 2(b), this image is the best result possible of be obtained and it was captured in a controlled lighting environment. The black region on the
The image of Figure 2(b) has about 14.3%, approximately the image presented in Figure 2(a).

Figure 2(b) corresponds to the skin region in the original image in Figure 2(a).

The image of Figure 2(b) has about 14.3%, approximately $\frac{1}{7}$, of white pixels, thus the area with greater concentration may be considered seven times greater than the other.

Supposing that the both concentrations are independent gaussians, but each one has different means, standard deviations and amplitudes, the distributions of background $G_b$ and hand $G_h$ are given respectively by

$$G_b(Cb, Cr) = \frac{A_b}{\sqrt{2\pi\sigma_b}} e^{-\frac{\|Cb, Cr - \mu_b\|^2}{2\sigma_b^2}},$$
(8)

and

$$G_h(Cb, Cr) = \frac{A_h}{\sqrt{2\pi\sigma_h}} e^{-\frac{\|Cb, Cr - \mu_h\|^2}{2\sigma_h^2}},$$
(9)

where $\mu_b = \begin{bmatrix} \mu_{Cb} \\ \mu_{Cr} \end{bmatrix}_{background}$ and $\mu_h = \begin{bmatrix} \mu_{Cb} \\ \mu_{Cr} \end{bmatrix}_{hand}$, $\sigma_b$, and $\sigma_h$, and $A_b$ and $A_h$ are the means (vector), standard deviations and amplitudes for $G_b(Cb, Cr)$ and $G_h(Cb, Cr)$, respectively. Using the separability property of independent random variables [15], the distribution may be rewritten as following

$$G(Cb, Cr) = G(Cb)G(Cr),$$
(10)

this property will be useful to specify the parameters of the reference’s histogram for each component.

5.2 Determination of the gaussians’ parameters

So, to determine the standard deviation and amplitude of each gaussian, the following information is necessary

$$Ampl \{G_h(Cb, Cr)\} = 10Ampl \{G_b(Cb, Cr)\},$$
(11)

and

$$Area \{G_h(Cb, Cr)\} = 7Area \{G_b(Cb, Cr)\}. $$
(12)

The Equations 11 and 12 may be rewritten as

$$G_h(\mu_h) = 10G_b(\mu_h),$$
(13)

and

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G_h(x, y) dxdy = 7 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G_b(x, y) dxdy,$$
(14)

respectively.

Substituting the Equations 8 and 9 in the Equations 13 and 14 obtains the following relations

$$A_h = 7A_b$$
and
$$\sigma_h = \frac{A_h}{10A_b} \sigma_b.$$  
(15)

Thus, any values according to the Equation 15 may be used, but computational limitations introduce some restrictions.

The work [19] presents the values for $\mu_b = \begin{bmatrix} \mu_{Cb} \\ \mu_{Cr} \end{bmatrix}_{background}$:

$$\mu_b = \begin{bmatrix} 114.872 \\ 148.875 \end{bmatrix},$$
(16)

and $\mu_h = \begin{bmatrix} \mu_{Cb} \\ \mu_{Cr} \end{bmatrix}_{background}$ is obtained doing $R = G = B = 255$ in the Equation 7:

$$\mu_h = \begin{bmatrix} 128 \\ 128 \end{bmatrix}. $$
(17)

The value of $\sigma_h$ used is $\sigma_h = 1.0e - 3$, thus $\sigma_b = 7.0e - 4$. The value of $A_b$ used is $A_b = 1$, thus $A_h = 7$.

The desired histogram is obtained with the adding $G_h(Cb, Cr) + G_h(Cb, Cr)$. Using the Equation 10, the desired histogram for each component is given by

$$p(x) = G_b(x) + G_h(x).$$
(18)

The obtained reference’s histograms are shown in Figure 3.

6. RESULTS

The images used in the tests and the results obtained with the skin segmentation algorithm are shown on the Figure 4,
those images show three gestures of the Libras, the Figure 4(a) shows the letter 'A', the Figure 4(c) shows the letter 'B' and the Figure 4(e) shows the letter 'G'. This figure presents images with bad segmentation with inadequate illumination, and it modifies the captured colors from the skin.

The original images were captured in a place with natural illumination and their backgrounds are white predominant.

The Figure 5 shows the histograms of the image from the Figure 4(a). According with this figure, may be observed that the Cb component is very concentrated, seeming like the image have only one object.

The enhanced histograms obtained after the application of the proposed method are shown on the Figure 6. These histograms have two distinct concentrations: the greater is the background and the smaller is the hand.

The enhanced histograms \( p_z(z) \) are not equal to reference, presented in Figure 3, because the algorithm undergoes influences of the digitalization. The enhanced image and segmentation results are shown in Figure 7. After the enhancement, the skin regions are in evidence, but some noises were introduced on the segmented image.

To evaluate numerically the improvement of segmentation, the original images are segmented manually, shown in Figure 8, resulting in a reference image.

The presented segmentation results are similar of others results found in the literature, some of them were obtained in lighting controlled environments [9, 6, 13, 18, 16].

The segmentation results using enhanced and original are
According to the results, the proposed enhancement improves, in average, 75.30% the segmentation results. The results motivates the use of the proposed method as a image enhancement to improve the efficiency of hand gesture recognition in different illumination conditions.

### 7. CONCLUSION

This work proposed a technique to correct high lighting images for skin segmentation using histogram matching. Analyzing the obtained results, the conclusion is that the segmentation provided much better results after the application of the histogram matching than before. The proposed method is very efficient, robust and simple to develop. The average gain in the studied images is 75.30%, compared with reference images.

The efficiency of the proposed method is limited to the environment because the gaussians' parameters are dependents of the scene and it inserts few noises doing that non-skin regions were labeled as skin. The inserted noises can be removed using simple filters, like morphological filters. The images segmented obtained were not filtered because the focus of this work is to present the histogram matching.
8. ACKNOWLEDGEMENTS
The authors would like to thank Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and the Laboratório de Engenharia de Sistemas de Computação (LESC) for technical and financial support.

Special thanks to Federação Nacional de Educação e Integração de Surdos (FENEIS - CE) and Associação de Pais e Amigos de Deficientes Auditivos (APADA - CE) for their generous help with this project.

9. REFERENCES