

Skin Color Detection Using PCA-based Color Representation

Sungyoung Kim and JaepilKo

*Dept. Computer Engineering, Kumoh National Institute of Techonogy, Daehak-ro 61,
Gumi, Gyeongbuk 730-701, Korea
{sykim, nonezero}@kumoh.ac.kr*

Abstract

Skin color detection is the process that classifies unknown colors into skin or non-skin classes. Skin color detection is preliminary step for facial or gesture analysis and can reduce search space for next higher level processing. In this paper, we show that proposed PCA-based color representation can give better performance than other frequently used color spaces such as XYZ and Luv. For skin detection, we use two classification models – histogram model and elliptical boundary model from skin and non-skin colors. The experimental results show the PCA-based color representation is more efficient than other color representation and the existing PCA-based representation based on the two classification models.

Keywords: *Skin Detection, PCA, Histogram, Boundary Elliptical Model, Chrominance, Luminance, Classification*

1. Introduction

The utilization of computer vision technology has been becoming more active than in previous years in real life. In particular, facial and gesture analysis are very common technologies in many application such as smart TV, video surveillance, bio-informatics, human-machine interface, and etc. As an example, you can control volume or turn channel on TV with your hands or body gesture instead of remote controller. It is also possible for you to play games with your game consoles by using just your body motion. Likewise, facial and gesture analysis is very useful for many research and application area.

Faces or hands detection is essential step for next analysis or recognition for the faces or hands. There are many excellent researches on face or hand detection but automatic skin detection plays a major role in this field [1, 2]. Skin detection can speed up intermediate or higher level processing by reducing search space. Pixel based method [3] is a very popular approach in skin detection. Skin detection can be achieved by skin classification or segmentation at pixel level that is to classify the pixels on input images into skin and non-skin classes. Pixel-based skin detection has long history and continues to be interesting research topic until now [4, 5]. Skin detection at pixel level is not trivial process and several main factors such as illumination, camera characteristics, ethnicity, individual characteristics like age and sex, and other additional factors as shadow must be considered [6, 7].

Two important aspects in pixel-based skin detection are the choice of color spaces used for skin modeling and skin models for classification [6, 8-11]. The choice of color spaces are first step in skin detection. There are various kinds of color spaces but RGB color space is primary space. RGB color space is not good choice for skin modeling but Jones and Rehg [13] compared the performance of histogram and mixture models in skin detection based on RGB color space. Other color spaces can be computed by linear or non-linear transformation from RGB color space. As skin colors have less variance in chrominance than luminance, usually

only chrominance components except luminance component are used for skin modeling. Normalized RGB, CIE XYZ, perceptual color spaces (HIS, HSV, HSL), TSL, orthogonal color spaces (YCbCr, YIQ, YUV), perceptual uniform color space (Lab, Luv) are often used. There have been some comparative studies [8-11] but no one is superior to others [12]. Chenaoua and Bouridane [14] tried to combine several color spaces as a substitute for choosing the best color space. They used input vectors of 24-dimension elements from seven color spaces and convert the vectors to 3-dimensional vector by using PCA. PCA was used to compute the representative colors but the approach required intensive memory on a computer.

As mentioned earlier, skin detection is a kind of two-class classification problem. Classification model for skin detection should be defined to decide a class of input colors. There exist three important models – histogram model, Gaussian mixture models, and elliptical boundary model. Jones and Rehg [13] compared the performance of histogram and mixture models. They computed skin and non-skin histogram and built a skin classifier using maximum likelihood. They also used 16 Gaussian for skin and non-skin classes. Elliptical boundary model [15] is a parametric model for color skin. Lee and Yoo compared the model with Gaussian mixture models with six Gaussians for skin colors.

This paper explores contributions of PCA on skin color detection and shows that PCA-based color representation can give better performance than other frequently used color spaces such as XYZ and Luv. Skin colors in RGB color space are transformed to the new mathematical space using PCA. We will show the PCA-based representation of skin and non-skin colors give excellent skin detection results compared to the existing color spaces. We use two classification models – histogram model and elliptical boundary model. Skin color and non-skin color histogram are computed for skin detection but elliptical boundary model is created using just skin colors. We will show elliptical boundary model combining skin color and non-skin color give good performance on skin color detection.

The remainder of the paper is organized as follows: Section 2 describes the superiority of PCA-based color representation for skin color detection. Sections 3 and 4 summarize skin detection process based on histogram model and elliptical boundary model, respectively. Experimental results in Section 5 shows the performance of the PCA-based color representation. Finally, we conclude in Section 6.

2. Effectiveness of PCA-based Color Representation on Skin Color Detection

2.1. PCA on Color Vectors

Principal component analysis (PCA) is a process to transform a set of data into a set of linearly uncorrelated components. The computed components from input color space are called as principal components. A pixel in an image has three values of typically R, G, and B and so can be expressed as a vector of 3-dimension. If the image is of size $W \times H$, there will be a total of WH 3-dimension vectors and a color vector is represented as $\mathbf{x} = [R \ G \ B]$. Let \mathbf{X} be a matrix of the vector population, then the mean vector and covariance matrix of the vector population \mathbf{X} are defined as Eq. (1) where $E\{\mathbf{x}\}$ is the expected value of \mathbf{x} .

$$\begin{aligned} \mathbf{m}_x &= E\{\mathbf{x}\} \\ \mathbf{C}_x &= E\{(\mathbf{x} - \mathbf{m}_x)(\mathbf{x} - \mathbf{m}_x)^T\} \end{aligned} \quad (1)$$

It is always possible to find a set of n eigenvectors because \mathbf{C}_x is real and symmetric. Let \mathbf{A} be a matrix whose rows are the eigenvectors of \mathbf{C}_x according to their eigenvalues in descending order. The matrix \mathbf{A} is a $k \times 3$ in size where k is the number of eigenvectors and

$k \leq n$. The \mathbf{x} 's are linearly transformed into vectors \mathbf{y} 's by Eq. (2). Vector \mathbf{x} is transformed to a lower dimensional vector \mathbf{y} of dimension k .

$$\mathbf{y} = \mathbf{A}(\mathbf{x} - \mathbf{m}_x) \quad (2)$$

In this paper we use two kinds of data set from the ECU [8] and Compaq [13] to gather skin and non-skin training samples. ECU dataset consists of 4,000 images where 1% images are from digital cameras and the rest are from web. Compaq dataset consists of 13,640 images from web. ECU dataset contains only skin images but Compaq dataset has 4,675 skin and 8,965 non-skin images.

Training samples on skin colors are extracted from skin images in ECU dataset. As ECU dataset contain just skin images, we extract training samples on non-skin colors from non-skin images in Compaq dataset. Although Compaq dataset contain skin images we use ECU skin images to compare the performance of skin detection against the method of Chenaoua and Bouridane [14] with same dataset.

The number of color samples of skin and non-skin are over a hundred million, respectively. Due to the huge number of vectors and memory limitation, it may be impossible to calculate a transform matrix \mathbf{A} for PCA. In such case we can use R-SVD method [16, 17]. The R-SVD algorithm can carry out the SVD of a larger matrix $\mathbf{X}^* = (\mathbf{X} | \mathbf{N})$, where \mathbf{N} is an $H \times 3$ matrix consisting of H additional color vectors. By using R-SVD, we can update the Eigen-basis incrementally.

2.2. Luminance and Chrominance Separation based on PCA

First channel in PCA-based color representation corresponds to eigenvector with the largest eigenvalue among three eigenvalues. Eigenvector with the largest eigenvalue gives good representation for luminance in input images [18]. Color spaces such as Luv, TSL, YCbCr, YIQ, and HSI are chosen to separate luminance and chrominance. PCA-based color representation can give more energy-concentrated luminance component than other color spaces. Figure 1 shows three channels (components) applying PCA to a color image using Eigen-basis on skin colors. As you can see in the figure, most of variance due to the illumination is concentrated on the first component. On the other hand there is little variance in luminance around face and hand as you can see in the second component. The same is true in the third component.

General color spaces such as Luv and HSI represent colors for general purpose but PCA-based color representation is useful for specific purpose. Let's compare PCA-based representation with HIS color space. We compute standard deviation of skin colors from the input image in figure 1. The standard deviations of three components (from largest to smallest eigenvalue) in PCA-based color representation are 0.4107, 0.0443, and 0.0160 respectively, whereas standard deviations are 0.2345, 0.0793, and 0.2197 in the order of I, H, S. The component of largest eigenvalue in PCA-based representation has larger variation than I, but the second component in PCA has smaller variance than H or S.



Figure 1. PCA-based Color Representation

The first image is the first component of PCA that has a largest eigenvalue, and the second and the third are the second and third components, respectively.

3. Skin Color Detection based on Histogram Model

Histogram is well-known statistical structures for representing of the distribution of color samples. Each bin of a histogram counts the number of occurrence of the bin but can be easily converted into a probability distribution by dividing total number of data. Histogram model is very useful given large amount of training data. Jones and Rehg [13] showed that histogram model gave good skin detection results compared to Gaussian mixture models on same data.

In this paper we create two histogram models for skin and non-skin colors. Histogram model is a non-parametric model which estimate color distribution from training dataset. Histogram model can be created easily and quickly from given training dataset and is independent on the distribution of colors. We create skin color model with ECU dataset and non-skin color model with Compaq dataset. These two dataset provide sufficiently large training images for skin and non-skin color histogram. As we use only chrominance components in color images, all histogram except PCA3 have two-dimensional structure. Histogram for PCA3 is 3-dimension and are computed using PCA-based representation that are transformed from 24-dimensional color vectors [14].

Figure 2 shows the histogram on chrominance components of PCA2, normalized rgb, Lu^*v^* and XYZ for skin colors. Histogram size is 256×256 . PCA2, normalized rgb, and Lu^*v^* give intensive distribution in chrominance components. PCA2 is PCA-based 2-dimensional color representation by applying PCA to skin colors. Figure 3 shows the histogram on chrominance components for non-skin colors. Histogram of non-skin colors shows more dense distribution near gray point than skin colors.

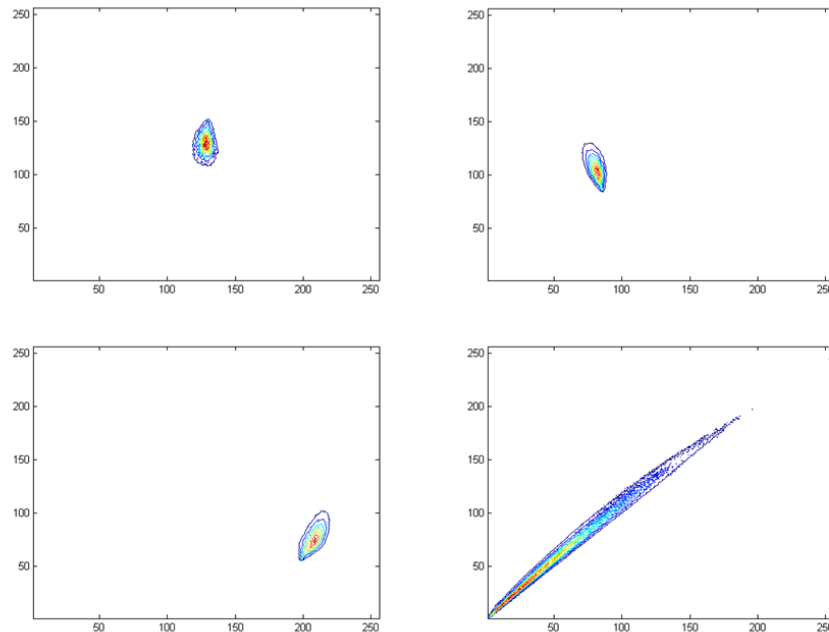


Figure 2. Contour Plots for Skin Chrominance Histogram with 256 Bins

The first one is for PCA2, the second is for normalized rg, the third is for u^*v^* and the last is for XY.

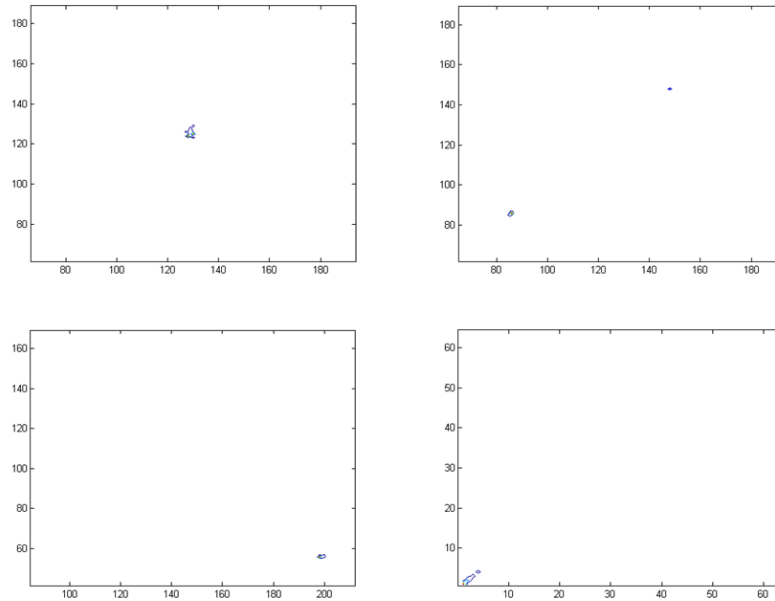


Figure 3. Contour Plots for Non-skin Chrominance Histogram with 256 Bins

The first one is for PCA2, the second is for normalized rg, the third is for u^*v^* and the last is for XY.

The probabilities of a given color that belongs to the skin and non-skin classes are computed by Eq. (3) given skin and non-skin color histogram model. S_{color} and N_{color} are the number of pixels with the given color in skin color histogram and non-skin color histogram, respectively. T_{skin} and $T_{non-skin}$ are the total number of pixels in the skin and non-skin histogram, respectively.

$$P(color|skin) = \frac{S_{color}}{T_{skin}} \quad (3)$$

$$P(color|non - skin) = \frac{N_{color}}{T_{non-skin}}$$

An unknown color is classified into a skin or non-skin class according to Eq. (4). The color is classified into skin when the ratio R is greater than a given threshold θ . Eq. (4) is the standard likelihood ratio. The threshold θ is between 0 and 1, and provides a trade-off between reducing false positive and increasing true positive.

$$R = \frac{P(color|skin)}{P(color|non - skin)} > \theta \quad (4)$$

4. Skin Color Detection based on Elliptical Boundary Model

Elliptical boundary model [15] is similar with single Gaussian model that use Mahalanobis metric but differs with the Gaussian model in using just distinctive chrominance value when creating classification model. Lee and Yoo [15] noted that skin colors have skewed distribution due to gray, and proposed the elliptical boundary model to address the problem. It gives better skin detection results against both single and mixture of Gaussians on the Compaq dataset.

The elliptical boundary model is defined as Eq. (5), and ψ and Λ are given by Eq. (6) where X_i 's are distinctive chrominance vectors of training data and $f(X_i)$ is the number of pixels with X_i . N is total number of pixels in training data and μ the mean of all chrominance vectors. A color X is classified into skin class if $\Phi(X) < \theta$. The threshold θ trades off true positives against false positives.

$$\Phi(X) = [X - \psi]^T \Lambda^{-1} [X - \psi] \quad (5)$$

$$\psi = \frac{1}{n} \sum_{i=1}^n X_i$$

$$\Lambda = \frac{1}{N} \sum_{i=1}^n f_i (X_i - \mu)(X_i - \mu)^T \quad (6)$$

The boundary model of skin colors gives good skin detection performance, but can be expanded with non-skin colors. If we create a boundary elliptical model with non-skin colors as well as skin colors, we can use the likelihood ratio in Eq. (7) to classify an unknown color into a skin or non-skin class. This ratio is very similar with the ratio in Eq. (4) of histogram model. The ratio in Eq. (7), however, has the form of that the numerator and denominator are inverted. This is because $P(\text{color}|\text{skin})$ and $P(\text{color}|\text{non-skin})$ have values of probabilities but $\Phi(\text{color}|\text{skin})$ and $\Phi(\text{color}|\text{non-skin})$ are metric of distance. $\Phi(\text{color}|\text{skin})$ and $\Phi(\text{color}|\text{non-skin})$ are distance of the color from the elliptical boundary model concerned with skin and non-skin colors, respectively.

$$R = \frac{\Phi(\text{color}|\text{non-skin})}{\Phi(\text{color}|\text{skin})} > \theta \quad (7)$$

5. Experimental Results

In this section we will show the performance of skin detection based on PCA-based color representation. Skin colors are extracted from ECU dataset [8] that consists of original and skin-segmented images. Non-skin colors are extracted from non-skin images in Compaq dataset [13]. The skin colors and non-skin colors in RGB color space are transformed to another mathematical space by using PCA and then histogram and elliptical boundary models are created based on the skin and non-skin colors, respectively.

We compare the skin color detection performance based on the PCA-based color representation against normalized rg, XY of XYZ, u^*v^* of Lu^*v^* and PCA3. All these color representation can be computed with linear or non-linear transformation from RGB color space. Lee and Yoo [15] showed that rg, XY, u^*v^* chrominance components gave best detection results among some color spaces, and so we compare the PCA-based color representation with the chrominance components in skin color detection. We call the PCA-

based color representation as PCA2. PCA3 is also PCA-based color representation but has a form of 3-dimensional vectors. The 3-dimensional vectors are transformed from 24-dimensional vectors by using PCA [14]. The 24-dimensional vector is made up of the following 24 components from several color spaces.

$$X = [R \ G \ B \ r \ g \ b \ n \ Y_{cbr} \ Cb \ Cr \ Y_{IQ} \ I \ Q \ H \ S \ V \ L \ a \ b \ T \ S \ rgrb \ YCB \ YCR]$$

rn, gn, bn are three components in normalized RGB where $rn=R/(R+G+B)$, $gn=G/(R+G+B)$, $bn=B/(R+G+B)$. If all of R, G and B have zero, then rn,gn and bn are set to $1/\sqrt{3}$. T and S are components of TSL that represent tint, saturation, and lightness. TSL color space was developed primarily for the purpose of face detection. rg and rb are computed as $(rn - gn)$ and $(rn - bn)$, respectively. YCB and YCR are $(Y - Cb)$ and $(Y - Cr)$, respectively.

We show the performance on skin detection using ROC curves. ROC means receiver operating characteristic and is used frequently for showing the performance of a binary classification. It plots true positive rate (TPR) against false positive rate (FPR). TPR is the fraction of true positives (TP) out of the total actual positives and FPR is the fraction of false positives (FP) out of the total actual negatives. You can see the meaning of TP and FP in Figure 4. Left table in Figure 4 is called confusion matrix that shows performance of classification.

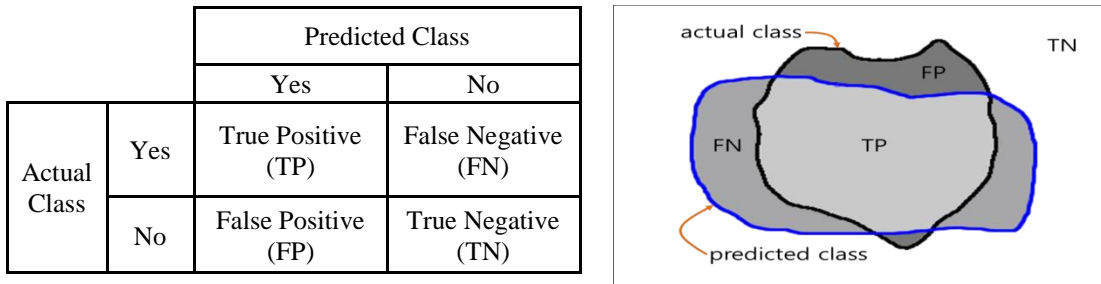
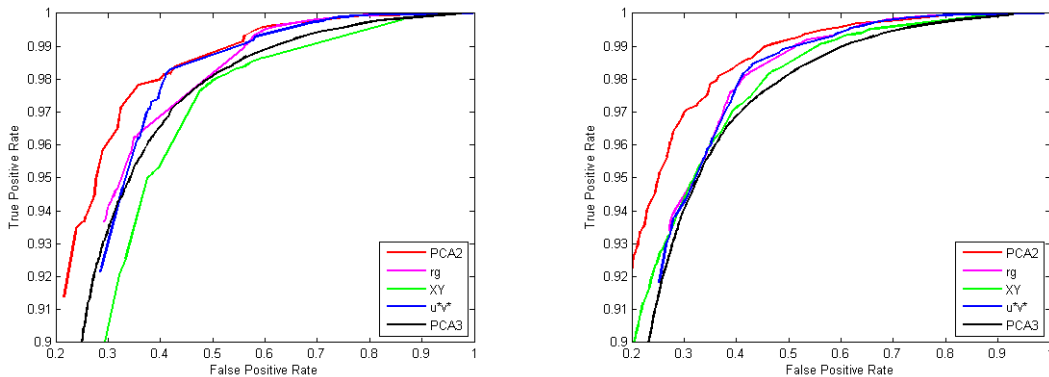


Figure 4. Confusion Matrix and its Visualization in skin Detection

Figure 5 shows ROC curves for the skin detection according to color space and histogram size. The ROC curves show effect on PCA2. PCA2 is superior in performance to other color spaces and especially to PCA3. The detection results are from test images.



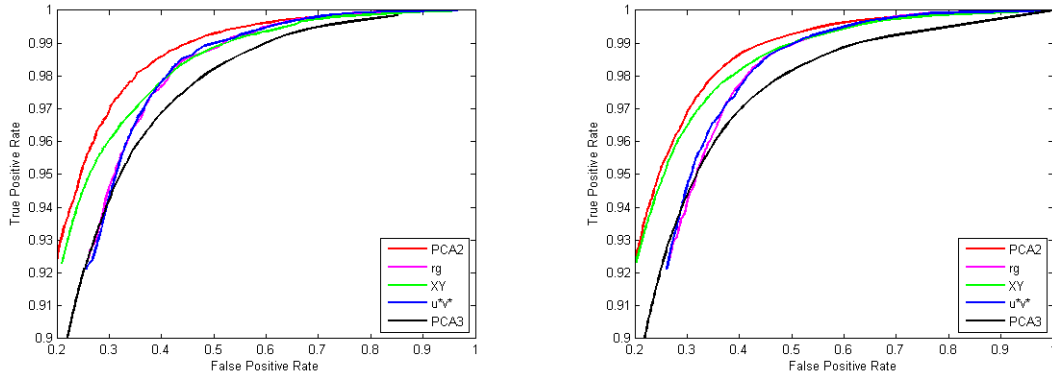


Figure 5. ROC Curves for the Skin Detection for Histogram Model as a Function of Color Space of PCA2, Normalized rg, XY, u^*v^* , and PCA3

The curve in left top is for histogram of 32×32 bins, right top is of 64×64 , left bottom is of 128×128 , and the last one is of 256×256 .

Figure 6 shows ROC curves of performance on skin detection changing histogram bin size. The bin size of histograms varies from 32 to 256 per a component. The ROC curves are created using test images. 64^2 , 128^2 and 256^2 sized histograms show subtle differences in detection performance but 256^2 histogram gives the best result among histogram models.

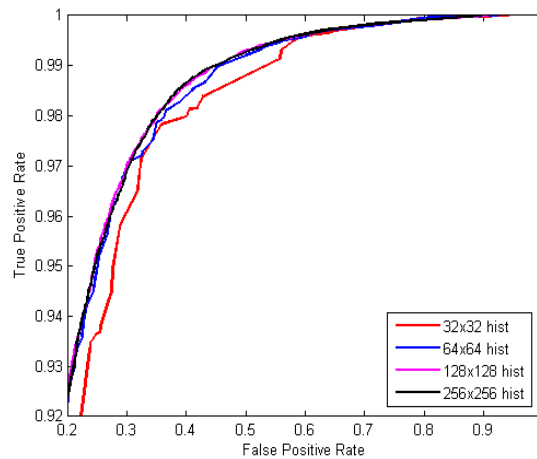


Figure 6. ROC Curves for the Skin Detection for Histogram Model as a Function of Histogram Size (from 32^2 to 256^2)

Figure 7 shows ROC curves for skin detection performance by using elliptical boundary model. The left curves in Figure 7 are computed with elliptical boundary model for just skin colors and the right ones are with both skin and non-skin colors. It is hard to select the best one in the left curves but it is trivial in the right curves. The overall TPR is improved in the right curves.

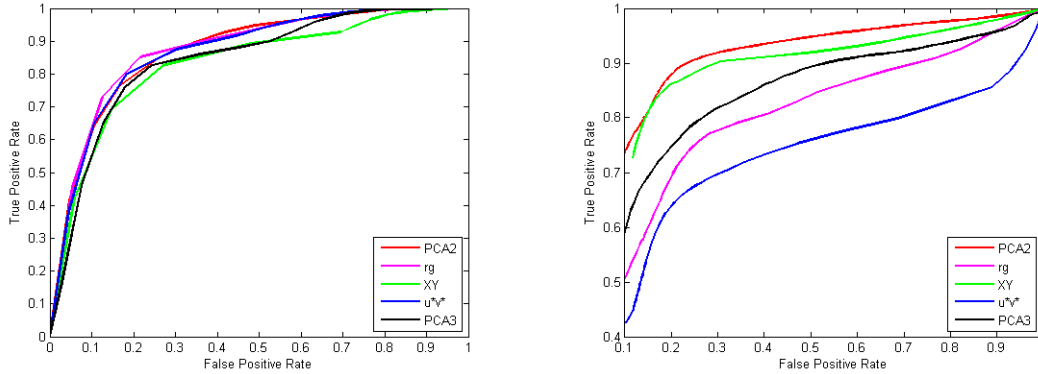


Figure 7. ROC Curves for the Skin Detection for Elliptical Boundary Model

The left curves are for detection ratio with just skin colors and the right ones are with both skin and non-skin colors.

Figure 8 shows ROC curves of performance on skin detection according to whether the use of non-skin or not. The elliptical boundary model on skin colors uses distance metric and we can get a curve that increase gradually by passing $(TPR, FPR) = (0, 0)$ changing a threshold from 0 to a maximum value. The metric in model using both skin and non-skin colors, however, is probability and a threshold is range from 0 to 1. TPR starts to increase from about 0.65 at a threshold of 1, and gets to 1.

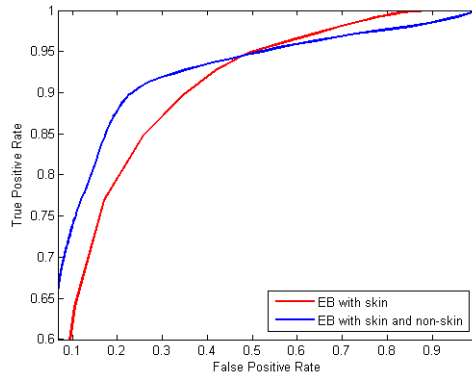


Figure 8. ROC Curves of Performance on Skin Detection According to whether the Use of Non-skin or Not

6. Conclusion

We explore contributions of PCA on skin color detection and shows that PCA-based color representation can give better performance than other frequently used color spaces such as XYZ and Luv. We found the 256^2 sized histogram and the elliptical boundary model with both skin and non-skin colors give good performance on skin color detection.

PCA on RGB colors is just linear transformation and can process quickly. PCA2 doesn't require intensive memory to compute as PCA3 does, but PCA2 shows superior detection results to PCA3. Skin color detection is preliminary step for facial or gesture analysis and can be used for applications for facial and gesture analysis that are very common technology in modern society.

Acknowledgements

This paper was supported by Research Fund, Kumoh National Institute of Technology.

References

- [1] J. Kovac, P. Peer, and F. Solina, "Human skin color clustering for face detection", EUROCON 2003, vol. 2, (2003), pp. 144-148.
- [2] M. Saaidia, S. Lelandais, and M. Ramdani, "Face Localization using Geometric and Skin Characteristics, International Journal of Signal Processing", Image Processing and Pattern Recognition, vol. 5, no. 3, (2012), pp. 9-26.
- [3] V. Vezhnevets, V. Sazonov, and A. Andreeva, "A Survey on Pixel-Based Skin Color Detection Techniques, Proceedings of the GraphiCon", (2003), pp. 85-92.
- [4] H.-M. Sun, "Skin detection for single images using dynamic skin color modeling", Pattern Recognition, vol. 43, (2010), pp. 1413-1420.
- [5] S. A. Naji, R. Zainuddin, and H. A. Jalab, "Skin segmentation based on multi pixel color clustering models", Digital Signal Processing, vol. 22, no. 6, (2012), pp. 933-940.
- [6] P. Kakumanu, S. Makrogiannis and N. Bourbakis, "A survey of skin-color modeling and detection methods", Pattern Recognition, vol. 40, no. 3, (2007), pp. 1106-1122.
- [7] A. M. Mharib, S. Mashohor, A. R. Ramli and M. H. Marhaban, "The Impact of Light Compensation on the Performance of Parametric Skin Detection Model", International Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 4, no. 4, (2011), pp. 29-37.
- [8] S. L. Phung, A. Bouzerdoum, and D. Chai, "Skin segmentation using color pixel classification: Analysis and comparison", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 1, (2005), pp. 148-154.
- [9] A. Albiol, L. Torres, and E. J. Delp, "Optimum color spaces for skin detection", Int'l Conf. on Image Processing, vol. 1, (2001), pp. 122-124
- [10] A. Elgammal, C. Muang and D. Hu, "Skin Detection - a Short Tutorial, Encyclopedia of Biometrics", (2009), Springer.
- [11] G. Gomez, "On selecting colour components for skin detection", Int'l Conf. on Patter Recognition, vol. 2, (2002), pp. 961-964.
- [12] M. C. Shin, K. I. Chang, and L. V. Tsap, "Does Color Space Transformation Make Any Difference on Skin Detection?", Workshop on Applications of Computer Vision, (2002), pp. 275-279.
- [13] M. J. Jones and J.M. Rehg, "Statistical color models with application to skin detection, Int'l Conf. on Computer Vision and Pattern Recognition", vol. 1, (1999), pp. 274-280.
- [14] K. Chenaoua and A. Bouridane, "PCA Based Choice of Representative Colors for Skin Detection, EUSIPCO - European Signal Processing Conference, (2005).
- [15] J. Y. Lee and S. I. Yoo, "An elliptical boundary model for skin color detection", Int'l Conf. on Imaging Science, Systems and Technology, (2002).
- [16] D. Ross, J. Lim and M. H. Yang, Adaptive Probabilistic Visual Tracking with Incremental Subspace Update, Proc. of the 8th European Conf. on Computer Vision, (2004), vol. 2, pp. 470-482.
- [17] L. Wang, L. Wang, Q. Zhuo, H. Xiao and W. Wang, "Adaptive Eigenbackground for Dynamic Background Modeling", Intelligent Computing in Signal Processing and Pattern Recognition, vol. 345, (2006), pp. 670-675.
- [18] L. Meylan and S. Ssstrunk, "High dynamic range image rendering using a Retinex-based Adaptive filter", IEEE Transactions on Image Processing, vol. 15, no. 9, (2006), pp. 2820-2830.

Authors

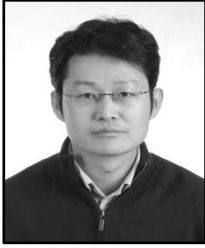


Sungyoung Kim

E-Mail: sykim@kumoh.ac.kr

Research Interest: Image processing, Pattern recognition, Multimedia information processing.

He received Ph. D. in computer engineering from Pusan National University, Korea in 2004. From 2004, he is a professor in dept. of computer engineering at Kumoh National Institute of Technology.



JaepilKo

E-Mail: nonezero@kumoh.ac.kr

Research Interest: Pattern recognition, Image processing

He received Ph. D. in computer science from Yonsei University, Korea in 2004. From 2004, he is a professor in dept. of computer engineering at Kumoh National Institute of Technology.

