Pertanika J. Sci. & Technol. 15(2): 131 - 137 (2007)
ISSN: 0128-7680

Pixel-Based Skin Colour Detection Techniques Evaluation

Ahmed M. Mharib[1], Mohammad Hamiruce Marhaban[2] and Abdul Rahman Ramli[1]

Departments of [1]Computer and Communication Systems Engineering
[2]Electrical and Electronic Engineering
Faculty of Engineering, Universiti Putra Malaysia,
43400 UPM, Serdang, Selangor, Malaysia
E-mail: ahmad_muhseen@yahoo.com

ABSTRACT

Skin detection has gained popularity and importance in the computer vision community. It is an essential step for important vision tasks such as the detection, tracking and recognition of face, segmentation of hand for gesture analysis, person identification, as well as video surveillance and filtering of objectionable web images. All these applications are based on the assumption that the regions of the human skin are already located. In the recent past, numerous techniques for skin colour modeling and recognition have been proposed. The aims of this paper are to compile the published pixel-based skin colour detection techniques to describe their key concepts and try to find out and summarize their advantages, disadvantages and characteristic features.

Keywords: Skin colour detection, computer vision community, pixel - based methods

INTRODUCTION

Skin segmentation is a computer vision process that aims to locate the skin regions in an input image. In the pixel-based skin detection methods, the task can be considered as a standard two-class classification problem; taking each pixel of the input image and producing binary output image that represents both the skin pixel and non-skin pixel. Methods that take pixels spatial relationship into account use the pixel-based method for one of the stages of their algorithms, so they are also at least partially dependant on the performance of the pixel-based methods (Vezhnevets and Andreeva, 2005; Vezhnevets *et al.*, 2003).

The focus will be on the pixel-based methods which in general can be classified into three categories (Gomez and Morales, 2002; Lee and Yoo, 2002; Terrillon *et al.*, 2000; Vezhnevets and Andreeva, 2005; Vezhnevets *et al.*, 2003; Zarit *et al.*, 1999):

- Explicitly defined skin region.
- Nonparametric skin distribution modeling.
- Parametric skin distribution modeling.

HUMAN SKIN COLOUR CHARACTERIZATION

The term "skin colour" is not a physical property of an object, but it is rather a perceptual phenomenon and therefore a subjective human concept. Colour is a perceptual representation of the surface reflectance of an object, as a result of the human eye's sensitivity to electromagnetic radiation wavelengths. Therefore, colour representation similar to the colour sensitivity of the human vision system should help to obtain high performance of the skin detection algorithm (Acharya and Ray, 2005).

Detecting human skin using colour as a feature has several problems. First, the colour representation of a skin obtained by a camera is influenced by many factors such as light and shadow. Second, different cameras produce significantly different colour values even for the same person under the same light conditions. In addition to these, human skin colours differ from one person to another (Jayaram *et al.*, 2004).

COLOUR SPACE REPRESENTATIONS

A wide variety of colour spaces have been applied to the problem of skin colour modeling. Different skin detection methods use different colour spaces. Albiolt *et al.* (2001) showed that for RGB, YCbCr and HSV colour spaces; there exists an optimum skin detector scheme such that the performance of all the skin detector schemes is the same and the separability of the skin and non-skin classes is independent of the colour space chosen.

Jayaram *et al.* (2004) presented an evaluation study of the effect of the pixel colour transformation from the RGB colour space to a non-RGB colour space and the dropping luminance component of the skin colour. The results revealed in the research can be summarized as follows:

1. The improvement in performance due to the colour space transformation was present, but not consistent.
2. The absence of the luminance component decreases performance.
3. The skin colour modeling has a greater impact than the colour space transformation.

EXPLICITLY DEFINED SKIN REGION

A method to build a skin classifier is to define explicitly (through a number of rules) the boundaries skin cluster in some colour spaces. The simplicity of this method has attracted many researchers (Garcia *et al.*, 1999; Gomez and Morales, 2002; Peer *et al.*, 2003; Wimmer and Radig, 2005). Ahlberg (1999) used a simple Discriminates for Face Localization and Facial Feature Extraction to extract skin regions in the image. This method uses the chrominance components of the YCbCr colour space, and the discriminate is as below:

$$\left.\begin{array}{c} 138 < Cr < 178 \\ AND \\ 200 < (Cb + 0.6Cr) < 215 \end{array}\right\} \Rightarrow Skin\,pixel \qquad (1)$$

Where *Cb* and *Cr* are the blue and red chroma components.

NON-PARAMETRIC SKIN DISTRIBUTION MODELLING

The non-parametric skin modelling methods are used to estimate the skin colour distribution from the training data without deriving an explicit model of the skin colour.

Normalized Lookup Table

There are two issues that must be addressed in building a colour histogram model; the choice of colour space and the size of the histogram which is measured by the number of bins per colour channel. The colour space is quantized into a number of bins, each of which corresponds to a particular range of colour component value pairs (in 2D case) or triads (in 3D case). These bins, forming 2D or 3D histograms are referred to as the lookup table (LUT). Each bin stores the occurring times of the particular colour in the training skin images. After training, the histogram counts are normalized, converting histogram values to discrete probability distribution (Vezhnevets and Andreeva, 2005).

$$Pskin\ (c) = \frac{skin\ (c)}{Norm} \tag{2}$$

Where *skin(c)* gives the value of the histogram bin, corresponding to colour vector *c* and *Norm* is the normalization coefficient (sum of all histogram bin values, or maximum bin value present). The normalized values of the lookup table bins constitute the likelihood that the corresponding colours will correspond to skin.

To perform skin detection, an image is first transformed into the used colour spaces of LUT. For each pixel in the image, the colour values index the normalized value in the LUT. If this value is greater than a threshold, the pixel is identified as skin. Otherwise, the pixel is considered to be non-skin (Zarit *et al.*, 1999).

Bayes' Classifier

In contrast with the lookup table method, the Bayesian method uses two colour histograms, one for the skin and the other for the non-skin pixels. The probability distribution value, *Pskin(c)*, is actually a conditional probability *P(c|skin)* "a probability of observing colour c, knowing that we see a skin pixel". A more appropriate measure for skin detection would be *P (skin|c)* "a probability of observing skin, given a concrete c colour value". To compute this probability, the Bayes' rule is used (Vezhnevets *et al.*, 2003):

$$Pskin\ (c) = \frac{P(c|skin)P(skin)}{P(c|skin)P(skin) + P(c|\overline{skin})P(c|\overline{skin})} \tag{3}$$

P(c|*skin*) and $P(c|\overline{skin})$ are directly computed from the skin and non-skin colour histograms equation [2]. The prior probabilities P(*skin*) and $P(c|\overline{skin})$ can also be estimated from the overall number of skin and non-skin samples in the training set. An inequality P(c|*skin*)≥ Θ, where Θ is a threshold value, can be used as a skin detection rule.

PARAMETRIC SKIN DISTRIBUTION MODELING

The parametric statistical approaches represent the skin-colour distribution in the parametric form, such as the Gaussian mode. However, the skin-colour distribution is usually multimodal and cannot adequately be represented as a single Gaussian in the colour space. Therefore, the use of a Gaussian mixture model has been proposed. Typically, the Expectation-Maximization (EM) algorithm is employed to fit and update these models based on the observed data (Phung *et al.*, 2005).

Single Gaussian Modeling (SGM)

A Gaussian or normal distribution is a symmetrical frequency distribution having a precise mathematical formula relating the mean and standard deviation of the samples. Gaussian distributions yield bell shaped frequency curves having a preponderance of values around the mean with progressively fewer observations as the curve extends outward. The probability distribution function (PDF) used to describe the probability of the variate to belong to the group of data which have Gaussian distribution.

A multi-variate Gaussian distribution is a specific probability distribution, which can be thought of as a generalization to higher dimensions of the one-dimensional normal distribution. Skin colour distribution can be modelled by a multi-variate Gaussian joint probability density function (PDF), defined as (Terrillon *et al.*, 2000):

$$P\left(x|skin\right)=\frac{1}{\sqrt{2\pi}\sqrt{\left|\sum_s\right|}}e^{-\frac{1}{2}((x-\mu_s)^T\sum_s^{-1}(x-\mu_s))} \tag{4}$$

Here, x is a colour vector, d is the dimension of colour vector and μ_s and $\sum_s$ are the distribution parameters (Mean vector and covariance matrix, respectively). The model parameters are estimated from the training data as follow:

$$\mu_s=\frac{1}{n}\sum_{j=1}^{N}x_j \qquad \sum_s=\frac{1}{n-1}\sum_{j=1}^{N}(x_j-\mu_s)(xj-\mu_s)^T \tag{5}$$

where n is the total number of skin colour samples.

Fig. 1 shows an example of the bivariate Single Gaussian distribution, T and S represent the colure vector components and the vertical axis represents the PDF.

A particular input pixel of colour vector is classified as skin if the probability density function satisfies the following relation:

$$P\ (x|skin)\geq\Omega \tag{6}$$

where Ω here is a threshold value.

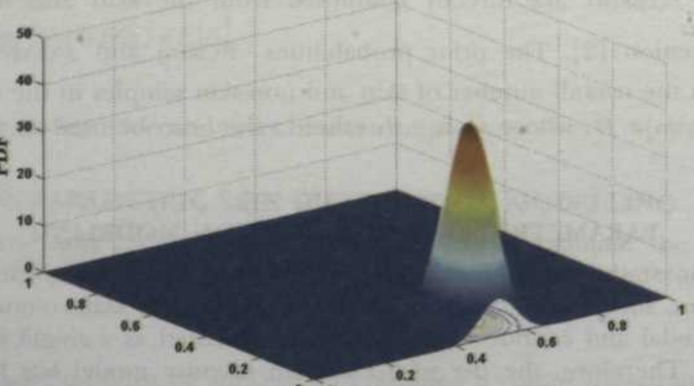

Fig. 1: Probability map of the bivariate Single Gaussian distribution

Gaussian Mixture Modeling (GMM)

A more sophisticated model, capable of describing complex-shaped distributions is the Mixture of Gaussians model (GMM). It is the generalization of the single Gaussian, the PDF in this case is:

$$P\ (x|skin) = \sum_{i=1}^{l} K_i \times P_i(x|skin) \tag{9}$$

Where l is the number of the mixture components K_i is the mixing parameters and $P(x|skin)$ are Gaussian PDF's, each with its own mean and covariance matrix. The model training is performed with a well-known iterative technique called the Expectation Maximization (EM) algorithm, which assumes that the number of components l to be known. The choice of the components number is important. The EM algorithm is run on the training data using a stopping criterion that checks if the change in log-likelihood between the two iterations is lesser than a threshold.

Greenspan *et al.* (2001) showed that a Gaussian mixture modelling of the colour space provides a robust representation to accommodate large colour variations. They presented a face segmentation model that used a Gaussian mixture modelling with k=2. They demonstrated this via the GMM, and the ability to provide better overall face segmentation results, as well as the ability to analyze the different regions within a face as belonging to different colour and lighting categories is achieved.

Fig. 2 shows an example of the bivariate Gaussian Mixture distribution, T and S represent the colure vector components and the vertical axis represents the PDF.

Elliptic Boundary Modeling

By examining the skin and non-skin distributions in several colour spaces, Phung *et al.* (2005) concluded that the skin colour cluster, being approximately elliptic in shape, was not well enough approximated by the single Gaussian model. Due to the asymmetry of the skin cluster with respect to its density peak, the usage of the symmetric Gaussian

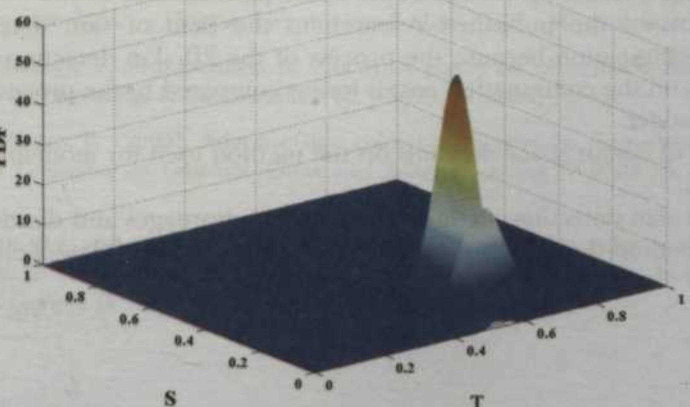

Fig. 2: Probability map of the bivariate Gaussian Mixture distribution

model leads to a high false positive rate. They proposed elliptical boundary model which is equally fast and simple in training and evaluation as the single Gaussian model.

COMPARATIVE EVALUATION FOR SKIN DETECTION METHODS

The main advantage of the methods that use explicitly defined skin cluster boundaries is the simplicity and intuitiveness of the classification rules. However, the difficulty with them is the need to find both good colorspace and adequate decision rules empirically.

The non-parametric methods are fast both in training and classification, independent to distribution shape and therefore to color space selection. However, they require much storage space and a representative training dataset. The parametric methods can also be fast, they have a useful ability to interpolate and generalize incomplete training data, they are expressed by a small number of parameters and need very little storage space. However, they can be slow (like mixture of Gaussians) in both training and work, and their performance depends strongly on the skin distribution shape. Besides, most parametric skin modeling methods ignore the non-skin color statistics.

Vezhnevets *et al.* (2003) and Phung *et al.* (2005) introduced a survey in skin detection methods. They both found that the Bayesian classifier with the histogram technique has higher classification rates compared to other tested classifiers, including piecewise linear and Gaussian classifiers. The Bayesian classifier with the histogram technique is feasible for the skin color pixel classification problem because the feature vector has a low dimension and a large training set can be collected. However, the Bayesian classifier requires significantly more memory compared to other classifiers. Parametric skin modeling methods (SGM, GMM and elliptic boundary model) are better suited for constructing classifiers in case of limited training and expected target data set. The generalization and interpolation ability of these methods makes it possible to construct a classifier with acceptable performance from incomplete training data.

CONCLUSION

From the previous review of the skin detection task, many important issues can be summarized as follows:

- Most of the recent studies with interest in the colour spaces and its relationship with the skin detection task, state that the elimination of the luminance component in the process of skin detection doesn't improve the performance of the model. Regardless of that, most of the published research in the field of skin detection used the luminance elimination because the process of the 2D skin detection model is easier and faster, and the computation cost is lowest compared to the process of the 3D skin detection model.
- The choice of colour space depends on the method used for modeling the skin color distribution.
- Each of the skin detection methods has its own advantages and disadvantages, so the choice of method depends on the application that the skin detection model is to be used.

REFERENCES

ACHARYA, T. and RAY, A. K. (2005). *Image Processing Principles and Applications.* New Jersey: A Wiley-Interscience Publication.

AHLBERG, J. (1999). *A system for face localization and facial feature extraction.* Technical report, Report no. LiTH-ISY-R2172. Linköping, Sweden: Linköping University.

ALBIOLT, A., TORRES, L. and DELP, E. J. (2001). Optimum color spaces for skin detection. In *Proceedings of IEEE International Conference on Image Processing* (*1*, 122–124). Thessaloniki, Greece.

GARCIA, C., ZIKOS, G. and TZIRITAS, G. (1999). Face detection in color images using wavelet packet analysis. *IEEE International Conference on Multimedia Computing and Systems, 1*, 703-708.

GOMEZ, G. and MORALES, E. F. (2002). Automatic feature construction and a simple rule induction algorithm for skin detection. In *Proceedings of the ICML Workshop on Machine Learning in Computer Vision* (p. 31-38). Sydney, Australia.

GREENSPAN, H., GOLDBERGER, J. and ESHET, I. (2001). Mixture model for face color modeling and segmentation. *Pattern Recognition Letters, 22*(14), 1525-1536.

JAYARAM, S., SCHMUGGE, S., SHIN, M. C. and TSAP, L. V. (2004). Effect of colorspace transformation, the illuminance component, and color modeling on skin detection. In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04)* (p. 813-818). Washington, DC, USA.

LEE, J. Y. and YOO, S. I. (2002). An elliptical boundary model for skin color detection. In *The 2002 International Conference on Imaging Science, Systems, and Technology.* Las Vegas, USA.

PEER, P., KOVAC, J. and SOLINA, F. (2003). Human skin colour clustering for face detection. In *EUROCON 2003 - International Conference on Computer as a Tool.* Ljubljana, Slovenia.

PHUNG, S. L., BOUZERDOUM, A. and CHAI, D. (2005). Skin segmentation using color pixel classification: analysis and comparison. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 27*(1), 148-154.

TERRILLON, J. C., SHIRAZI, M. N., FUKAMACHI, H. and AKAMATSU, S. (2000). Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images. In *Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition* (p. 54-61). Grenoble, France.

VEZHNEVETS, V. and ANDREEVA, A. (2005). *A comparative assessment of pixel-based skin detection methods.* Technical report. Moscow, Russia: Graphics and Media Lab., Moscow State University.

VEZHNEVETS, V., SAZONOV V. and ANDREEVA A. (2003). A survey on pixel-based skin color detection techniques. In *Proceedings of the Graphicon* (p. 85-92). Moscow, Russia.

WIMMER, M. and RADIG, B. (2005). Adaptive skin color classificator. In *Proceedings of the first ICGST International Conference on Graphics, Vision and Image Processing GVIP-05* (p. 324-327). Cairo, Egypt.

ZARIT, B.D., SUPER, B.J. and QUEK, F.K.H. (1999). Comparison of five color models in skin pixel classification. In *Proceedings of IEEE International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems* (p. 58-63). Corfu, Greece.