A Novel Adaptive Colour Segmentation Algorithm and Its Application to Skin Detection [+]

Yuchun Fang Tieniu Tan
National Laboratory of Pattern Recognition (NLPR),
Institute of Automation, Chinese Academy of Sciences
P. O. Box 2728, Beijing 100080, P. R. China
{ycfang, tnt}@nlpr.ia.ac.cn

Abstract

In this article, we bring forward a novel model-based region segmentation algorithm in colour images and use it to segment the skin-colour area from the background area adaptively. First, a quantitative colour space selection scheme is used to select a suitable colour representation manner. A two-fuzzy-set colour distribution model is adopted to depict the object area colour distribution in the selected colour space. Segmentation is performed by use of the α-cuts of the fuzzy sets. The value of α is adaptively decided to different images. This method is successfully used to segment skin colour area in arbitrary colour images with complex background and varying lighting conditions.

1 Introduction

Image segmentation is a hard and fundamental task in image processing and computer vision. Because of the diversity of the purpose that various image processing tasks need to fulfil, there exists no optimum method to achieve successful segmentation in all images.

With the development of image processing and analysis techniques, more and more work tends to use colour as a cue to process images. Although shape information is a more reliable property of objects, there do exist many natural or artificial objects for which colour can be an identifying feature [1]. In our work, we use colour cue to segment object areas out of the background.

According to the colour mechanism of human vision, a colour is quantitatively represented as a three-dimensional vector. So colour is a kind of high dimensional information, the traditional segmentation scheme based on one-dimensional intensity cannot be directly generalised to colour image segmentation. To process high

[+] This work is funded by research grants from the NSFC (Grant No. 59825105), the 863 Program (Grant No. 863-306-ZT-06-06-5 and 863-317-01-10-99) and the Chinese Academy of Sciences.

dimensional information of colour, one way is to decrease the dimension. Some work adopted general grey image processing methods in a one-dimensional colour space generated by selecting one component of the high dimensional colour space. Dai and Nakano detected faces with the ***I*** component of the ***YIQ*** colour space [2]. Lee et al. detected facial features with the ***H*** component of the ***HSV*** colour space [3]. It is apparent that this kind of method cannot adequately take advantage of the colour information. A more frequently adopted scheme uses colour distribution models (CDM) to perform region segmentation with colour information [4-10]. Histogram based CDM was adopted in reference [4] and [5] to detect skin area in a colour image. In reference [6] and [7], the CDM was constructed based on the fuzzy sets. A great deal of work is based on Gaussian models [8-10]. Colour models can make use of the colour information more fully, but such models cannot adapt to colour variations caused by the change of intensity and the sophisticated surface of objects. Furthermore, modeling with certain probability density functions such as Gaussian functions is based on a distribution hypothesis, which may not be realistic.

In our work, we also use a model-based scheme, and we perform segmentation dynamically to adapt to colour variations. In Section 2, we first introduce a quantitative colour space selection method we adopted in order to improve the segmentation performance. The construction of the CDM is introduced in Section 3. In Section 4, how to realise adaptive segmentation is described. The experimental results of colour space selection and adaptive region segmentation are given in Section 5.

2 Colour Space Selection

To process colour images, the first task is to select a suitable quantitative representation of colour, that is to say, to select a suitable colour space. Many different schemes have been adopted in previous work to represent colour information. Since some schemes take use of colour information through dimension-decreasing method, the dimension of a colour space may be one, two or three. Because different colour space has different advantages in representing colour, it is necessary to select a relatively better colour space according to the requirement of the task.

In our work of segmenting an area of certain colour out of the background, we construct a measure to evaluate the performance of various colour spaces, and then select the best one according to the evaluation measure.

Let O denote the colour space. Let A and B denote the object area and the background area respectively. Let $a_i^{(o)}$ and $b_j^{(o)}$ represent the colour vector of the i-th pixel of A and the j-th pixel of B in colour space O respectively, where i=1...m, j=1...n. The evaluation function is defined as follows:

$$D_o(A,B) = \frac{d(\bar{a},\bar{b})}{\sigma_o(A)+\sigma_o(B)}, \tag{1}$$

where $\bar{a}$, $\bar{b}$ are the mean of A and B respectively, and $d(\cdot,\cdot)$ denotes a distance measure, $\sigma_o(A)$ and $\sigma_o(B)$ are calculated as follows: if O is one-dimensional, they are the variance of A and B; if O is two-dimensional or three-dimensional, they are the trace of the covariance matrix of A and B.

We can see that the numerator of Equation (1) denotes the overall colour difference of area *A* and area *B*, the denominator of Equation (1) denotes the variations of the colour distribution in area *A* and area *B*. The larger the numerator and the smaller the denominator, the more predominant the difference between area *A* and *B*. In other words, the greater the value of $D_o(A,B)$, the easier area *A* can be discriminated from area *B* according to their colour representation in colour space *O*. From the above analysis we can see that the evaluation criterion defined in (1) is with similar idea in feature selection. So it is reasonable to use this formulation to select a colour space suitable for representing colour in order to detect object areas in a colour image.

3 Colour Distribution Model Construction

Our task is to segment objects with certain colours out of the background, so it is the difference between object colours and other colours that we make use of for segmentation. In such circumstance, taking good use of colour information means finding a good way to represent the difference. Like much of the other work [4,6,8-10], we use CDM to reflect such difference.

We model colour distribution in the selected colour space based on fuzzy sets theory. A fuzzy set *F* is depicted by an ordinary set *X* and a membership function μ_F on it. The range of μ_F is [0,1]. The function value $\mu_F(x)$ is used to denote the degree how the element $x \in X$ possesses certain property defined by *F*. The bigger the value of $\mu_F(x)$, the more likely the element *x* possesses the property of *F*. For example, we can use the membership function value of a colour to depict the degree how it belongs to certain objects' colours. We can see that, in a fuzzy set, this single membership value $\mu_F(x)$ is used to represent both the evidence how *x* possesses the certain property of *F* and the evidence how it does not possess such property without indicating how much there is of each. This single value cannot depict the accuracy of the membership value [11]. In order to improve this deficiency, Gau and Buehrer came up with the vague set theory based on two membership functions, a truth membership function and a false membership function [11]. They reflect the low bounds of the membership degree how an element belongs or does not belong to *F*. In our work, unlike the fuzzy model in the work of others such as Wu et al. [6], which used only one fuzzy set to build up CDM, we construct a two-membership-function model to depict the colour distribution. A truth membership function is used to denote the degree how a colour belongs to the object colours, and a false membership function denotes the degree how a colour does not belong to the object colours. These two fuzzy sets work together conceiving the fuzzy the CDM in our work.

To construct a fuzzy set, the key problem is to construct the membership functions. Let *O* denote the set of all colours in the colour space, and *x* denote an arbitrary element in this colour space. Our fuzzy CDM is constructed by creating two discrete membership functions, a truth membership function $t_o(x)$ and a false membership function $f_o(x)$, according to the statistic result of a large amount of object colour and non-object colour samples (positive and negative) respectively (where each sample is a pixel).

Let *L* and *M* denote the number of positive and negative samples respectively. We first construct the distribution histogram H^P of the positive samples and H^N of the negative samples in *O*. We divide space *O* into *k* bins, each of which can be used to represent a discrete colour in *O*. For the positive sample histogram H^P, the number of samples falling in the *i-th* bin is expressed as h_i^P, *i=1...k*, and for the negative sample histogram H^N, it is expressed as h_i^N, *i=1...k*. Then it is clear that $L = \sum_{i=1}^{k} h_i^P$ and $M = \sum_{i=1}^{k} h_i^N$. The more the positive samples falls in the bin, the more likely the colour represented by the bin is of object colour, and vice versa.

Based on H^P and H^N, we construct the discrete membership functions on *O* as follows:

$$\begin{cases} t_O(i) = \dfrac{\sum_{h_j^P \le h_i^P} h_j^P - h_i^P}{L} \\ f_O(i) = \dfrac{\sum_{h_j^N \le h_i^N} h_j^N - h_i^N}{M} \end{cases}, \tag{2}$$

where *i=1...k*, and each represents a discrete colour in *O*. The right side of each formulation in Equation (2) is a ratio between the sum of bin values no larger than the value of the *i*-th bin (excluding the *i*-th bin) and the number of total samples. Apparently, its value lies in interval [0,1] and its monotony is accordant with the monotonicity of the histogram. So the two formulations in Equation (2) can be used to depict how a colour belongs to object colour and how it does not belong to object colour respectively. It is in accordance with the definition of fuzzy sets. That is to say, The larger the value of $t_o(i)$ and the smaller the value of $f_o(i)$, the more likely the *i*-th colour in *O* belongs to object colours. So it is reasonable to use Equation (2) as a CDM for object area detection.

In Equation (2), the normalization is not performed by each bin value dividing the biggest bin value or by each bin value dividing the total number of samples. The former normalization will produce membership degree value 1 that means the colour is definitely the object colour, which is not reasonable in practice. The latter normalization may generate very small membership degree values, so it cannot provide an effective colour distribution representation. However the normalization scheme defined in Equation (2) produces no 1-value or very small values and it is therefore more consistent with our intuition.

4 Adaptive Segmentation

In a fuzzy set *F*, the membership function value reflects the degree how an element possesses certain property, and a α*-cut* of the fuzzy set is a subset whose elements' membership values are more than α. We also use α*-cuts* to our fuzzy CDM to partition the colour space as object colour subset and non-object colour subsets.

A truth membership function and a false membership function compose our CDM, so we use two thresholds α_t and α_f to derive two subsets S_t and S_f as follows:

$$\begin{cases} S_t = \{x \mid t_O(x) \ge \alpha_t \ \& \ x \in O\} \\ S_f = \{x \mid f_o(x) \ge \alpha_f \ \& \ x \in O\} \end{cases}. \tag{3}$$

Then the object colour subset S in O is calculated by

$$S = S_t \cap (O - S_f)\ . \tag{4}$$

If the colour of a pixel in an image belongs to S, this pixel is regarded as the object area in the image.

However, because of variations in lighting conditions and the complexity of the object surface, a fixed threshold is not suitable to all circumstances. In Jones and Rehg's work, they used a very large number of samples to construct a colour histogram model [4]. Large number of samples can give more sufficient information about object colour distribution in a given colour space, and it can improve the reliability of the model. But segmentation results are also decided by how we use the model. In different images, object area colours occupy different subsets of a colour space. And colour distribution of the object areas in a certain image is often not the same as that of a statistical distribution of all object colours in the selected colour space. Segmentation with a good CDM but a fixed segmentation criterion (e.g. thresholds) to all images also cannot always provide good segmentation result. In our work, we use adaptive thresholds to realise dynamic segmentation. The values of the two thresholds α_t and α_f are not fixed, they are decided by the average luminance L of the image and the luminance distribution D of the given image.

In region detection tasks, the object areas as the main parts of an image often have relatively higher luminance when compared with the situation they are not the main parts of the images. When the whole luminance of the image is lower down, the luminance of the object part is also lower down. Moreover, in an image of low average luminance but high luminance variance, the object areas will be of relatively high luminance variance, and their colours will possess a relatively large area in the colour space. This assumption of the relationship between the image luminance, image luminance variance and the object area colours is reasonable in common sense. Our threshold selection scheme is constructed according to such an assumption. We use the average luminance of all pixels of the image to denote the image luminance L and use the quotient between L and the standard deviation of the luminance of all the pixels in the image to denote luminance distribution D. The final threshold is decided by Equation (5).

$$\alpha(L,D) = b + \frac{w_1}{1+e^{\frac{L-\theta_1}{\theta_2}}} + \frac{w_2}{1+e^{\frac{D-\theta_3}{\theta_4}}}\ , \tag{5}$$

where b is a bias, w_1 and w_2 are weights, and θ_1, θ_2, θ_3, θ_4 are control parameters.

In Equation (5), the threshold is modeled as a linear weighted sum of two sigmoid functions with a bias value b. One sigmoid function is with regard to the image luminance; the other is with regard to the image luminance distribution. Here, we adopt sigmoid functions to transform the values in a large interval into a relatively smaller one nonlinearly. In order to obtain appropriate thresholds, w_1 and w_2 are used to balance the impacts of image luminance and luminance distribution on thresholds; b, θ_1, θ_2, θ_3 and θ_4 are adjusted to obtain values of α_t and α_f in a sub-interval of [0,1]; θ_1, θ_2, θ_3 and θ_4 also control the changing of thresholds with regard to image luminance and luminance distribution. For example, to calculate the value of α_t , we set θ_1 and θ_2 so that it is increasing with the increase of *L*. This is accordant with our assumption.

5 Implementation and Experimental Results

Detecting and segmenting object regions with certain colours in an image can be used to perform object detection, which is necessary in many image understanding tasks. We use our region segmentation scheme in detecting skin colour areas in an image and the following summarises our experiment results.

5.1 Colour Space Selection

According to the colour space selection scheme presented in Section 2, we performed colour space selection on the following 8 frequently adopted colour spaces: ***RGB***, ***RG*** plane in ***RGB***, ***rg***, ***HS*** plane in ***HSV***, ***H*** coordinate of ***HSV***, ***S*** coordinate of ***HSV***, ***I*** coordinate of ***YIQ*** and a colour space transformed from ***RGB*** by **KL** transformation [12]. To test the performance of the above eight colour spaces, we selected 120 images containing people of different age, gender and race. These images are of different luminance. In these 120 images, we randomly selected 18,000 skin colour pixels (150 for each image) and 24,000 background pixels (200 for each image) as samples to denote object area and background area respectively. We calculated the evaluation value of each image in the above 8 colour spaces according to Equation (1). The statistical result is summarised in Table 1.

From Table 1 we can see that among the 120 testing images, 117 images have the greatest evaluation value in the rg space than in others. So according to our evaluation function, we selected ***rg*** space to perform skin area detection tasks. Figure 1 is a distribution histogram of skin colour samples in rg space, from which we also can see the skin colours in the ***rg*** space is centralised. Our result agreed with that of Terrillon et al. at this point [8], but his conclusion is drawn just from the observation of sample distribution contour images in two colour spaces, whereas ours is drawn from a reasonable quantitative evaluation.

	RGB	**RG (RGB)**	**rg**	**I (YIQ)**	**KL**	**H (HSV)**	**S (HSV)**	**HS (HSV)**
Number	0	0	117	3	0	0	0	0
Sum	46.40	44.21	714.68	142.23	58.33	62.64	68.25	62.99
Average	0.39	0.37	5.96	1.19	0.49	0.52	0.57	0.52

Table 1: Comparison of 8 Colour Spaces

Note: the first row is the number of images that have the greatest evaluation values in different colour spaces, the second line and the third row are the sum and average of the evaluation values in different colour spaces respectively.

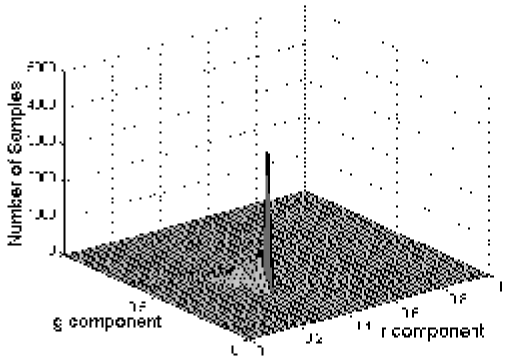

Figure 1. Distribution Histogram of Skin Colour Samples in rg Space

5.2 Segmentation Based on Fuzzy CDM

To construct our fuzzy CDM, we have selected more than 28,000 skin area colour pixels and 50,000 non-skin area colour pixels to construct the truth membership function and the false membership function. These sample pixels were randomly extracted from over 180 images that were scanned by a scanner or grabbed by a CCD camera. We tested our algorithm with images including those which were not used in fuzzy CDM construction. These images include both simple background images and complex background images taken both indoors and outdoors. Examples of segmentation results are shown in Figure 2. The blue areas denote the non-skin-colour area in the images and the skin colour areas are retained.

We compared the segmentation scheme using just a single false or truth membership function with the scheme that combines the two fuzzy membership functions. We also compared the effect of adaptive segmentation with that of segmenting with fixed thresholds. As illustrated in Figure 2, in the upmost row are the original images. They are ordered according to the image luminance. From left to right, the luminance value of the image is increasing. The segmentation results using only a false membership function or a truth membership function is presented in the second row and the third row respectively. The fourth row shows the processing results gained by combining two membership functions together but with fixed thresholds and in the last row are the segmentation results of adaptive thresholds and combination of two fuzzy membership functions. It is apparent that each of the two fuzzy membership functions works as a filter to eliminate certain non-object pixels, thus the two-fuzzy-set CDM greatly improves the segmentation result. We can also see that when the image luminance is too large or too small, an adaptive threshold can avoid missing object colour pixels and wrongly including non-object colour pixels.

We have also tested the same algorithm both in the ***rg*** space and the ***HS*** space. Even under a fixed threshold, the segmentation scheme can gain relatively better segmentation result in the ***rg*** space, but it is difficult to select a threshold to segment.

In summary, the results given above demonstrate that the algorithm described in this paper performs effectively in region segmentation tasks.

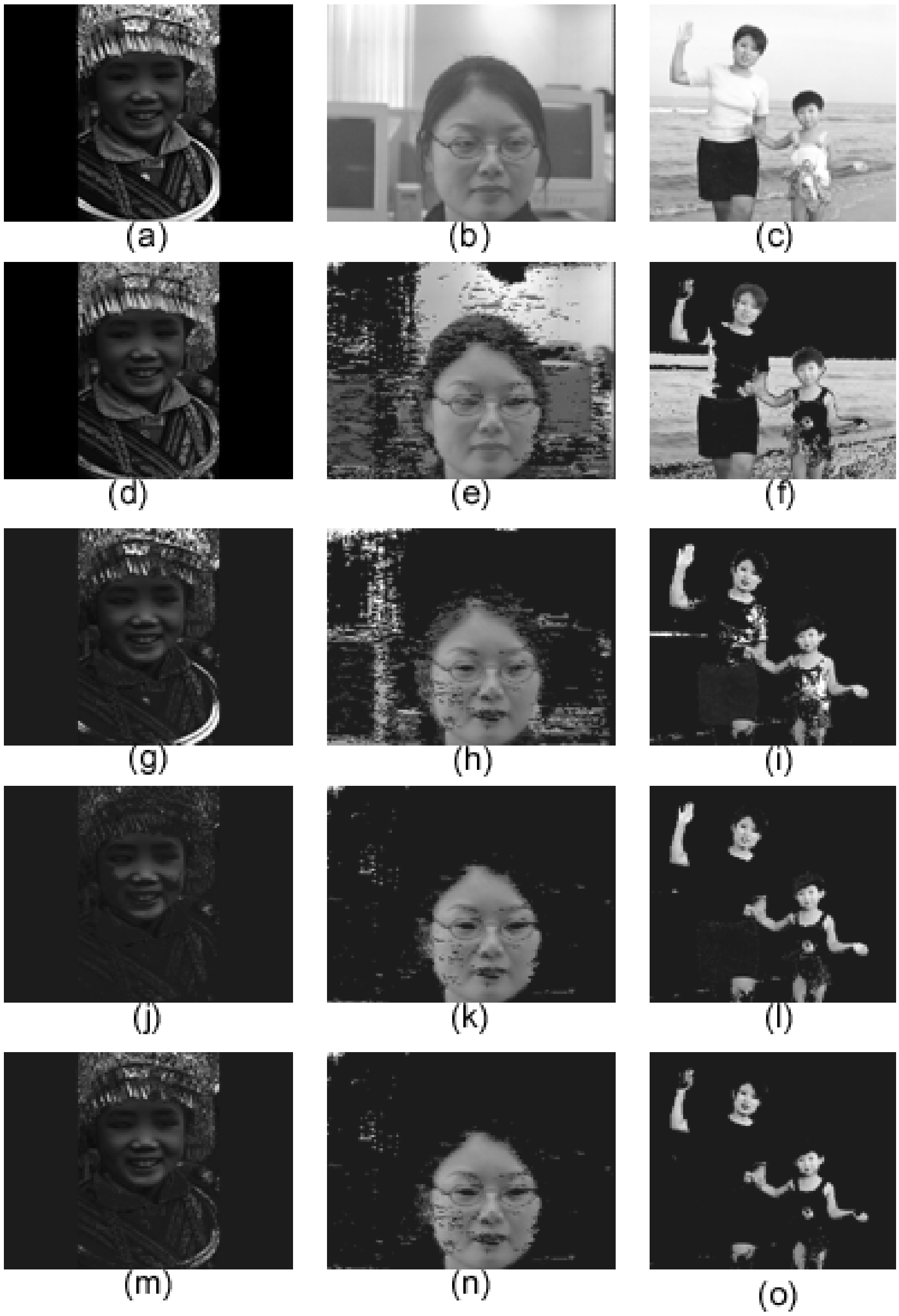

Figure 2. Comparison of different CDM based segmentation schemes

Note: (a)-(c), the original images; (d)-(f), segmentation results with false membership functions; (g)-(i), segmentation results with truth membership functions; (j)-(l), segmentation results by combining two membership function but using fixed thresholds; (m)-(o), segmentation results by combining two membership function and selecting thresholds adaptively.

6 Conclusion and Future Work

In this article, an effective quantitative evaluation method is brought forward to select a colour space. A new colour distribution model composed of two fuzzy membership functions is used to segment the skin region. It works like combining two models compensating with each other. We used an adaptive threshold selection method to cope with luminance variation. Because the luminance variance in an image is often too complex to be predicted, the segmentation results will probably be improved by using local luminance distribution of a single pixel instead of luminance distribution inside a whole image to adjust the threshold. In the experiment, the number of pixels used to build up our colour distribution model is relatively small. More data will improve the colour distribution model and thus improve the segmentation results. Further more, the algorithm in Section 1 to Section 4 is not limited to skin-colour region detection, it can also be applied to segment other objects that have certain colour.

References

[1] Swain, M. J., Ballard, D. H., Colour indexing, *International Journal of Computer Vision*, 1991, 7:1, pp.11-32.

[2] Dai, Y., Nakano, Y., Face-texture model based on SGLD and its application in face detection in a colour scene, *Pattern Recognition,* 1996, Vol.29, No.6, pp.1007-1017.

[3] Lee, C. H., Kim, J. S., Park, K. H., Automatic human face location in a complex background using motion and colour information, *Pattern Recognition*, 1996, Vol.29, No.11, pp.1877-1889.

[4] Jones, M. J., Rehg, J. M., Statistical colour model with application to skin detection, IEEE Int. Conf. on Computer Vision and Pattern Recognition, 1999. 6.

[5] Birchfield, S., Elliptical head tracking using intensity gradients and colour histograms, IEEE Int. Conf. on Computer Vision and Pattern Recognition, 1998.1.

[6] Wu, H., Chen, Q., Yachida, M., Detecting human face in colour image, IEEE Int. Conf. on Systems, Man and Cybernetics, 1996, Vol.3, pp.2232-2237.

[7] Ito, N. et al., The Combination of edge detection and region extraction in nonparametric colour image segmentation, *Information Sciences*, 1996, Vol. 92, pp277-94.

[8] Terrillon, J., David, M., Akamatsu, S., Detection of human face in complex scene images by use of a skin colour model and of invariant Fourier-Mellin moments, 14th Int. Conf. on Pattern Recognition, 1998, Vol.2, pp.1350-1355.

[9] Kim, S. H., et al., Object oriented face detection using range and colour information, Proc. of the 3rd IEEE International Conference on Automatic Face and Gesture Recognition, 1998, pp.76-81.

[10] Raja, Y., McKenna, S. J., Gong, S., Segmentation and tracking using colour mixture models，In Proc. of the 3rd Asian Conf. on Computer Vision, 1998, Vol.1, pp.607-14.

[11] Gau, W. L., Buehrer, D. J., Vague sets, *IEEE Transactions on Systems, Man and Cybernetics*, 1993, Vol.23, No.2, pp.610-14;

[12] Ohta, Y., Kanade, T., Sakai, T., Colour information for region segmentation, *Computer Graphics and Image Processing*, 1980, Vol.13, pp.222-41.