VOL. 10, NO. 16, SEPTEMBER 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

ADAPTIVE MEAN SHIFT FOR SKIN IMAGE SEGMENTATION

Kyoung-Mi Lee
Department of Computer Science, Duksung Women's University, Seoul, Korea
E-Mail: kmlee@duksung.ac.kr

ABSTRACT

The mean-shift clustering is an efficient technique for color image segmentation by dividing an image into homogeneous regions. The main drawback of mean-shift clustering is to use a fixed scale, which directly determines to use a fixed homogeneity. Since each region could have different homogeneity, using a fixed scale has a problem to segment well. To resolve this problem, we incorporate multi-resolution search by providing different scales to regions. The proposed algorithm starts initially at lowest resolution first and then proceeds to higher resolution where the search results are only refined. The proposed algorithm is applied to skin color segmentation.

Keywords: mean shift, coarse-to-fine search, skin detection, image segmentation.

INTRODUCTION

Image segmentation is the first essential and important step of image analysis and understanding (Pal and Pal, 1993). The goal of segmentation is to identify and group regions or features sharing similar characteristics together. There are several segmentation algorithms using thresholding, edge detection, region detection, or any combination of these techniques. The output of the image segmentation is usually a set of regions that cover the entire image, or a set of boundaries extracted from the image.

Mean-shift clustering is a nonparametric statistical method for seeking the nearest mode of a point sample distribution in its search window (Comaniciu and Peer, 2002). The effectiveness of mean-shift clustering has been demonstrated in computer vision such as region-based color image segmentation. Mean-shift segmentation groups homogeneous pixels into regions, and identifies the regions with unique modes by mapping each pixel to a mode using a convergent, iterative process. The advantages of mean-shift segmentation include flexibility in choosing the number of regions and noise robustness.

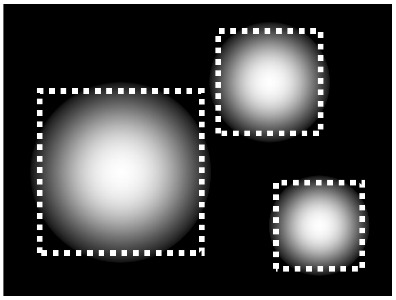


Figure-1. Variable mean-shift kernel scales

However, one of limitations of the mean-shift algorithm is to use a fixed scale which directly determines the size of the search window. Naturally, each pixel in an image could have different scale. Figure-1 shows a good single and global scale cannot segment well and thus variable kernel scales are needed for good segmentation. For example, skin color segmentation in a natural image accommodates an important challenge, varying illumination conditions. Depending on an amount of illumination on each pixel, the size of the search window in mean-shift should be various to determine appropriate homogeneity of regions. Some researchers have been studied by using adaptive bandwidth (Mohammadi and Mahzoun, 2012), using a difference of Gaussian kernel (Collins, 2003), or using a regression-based model (Fricker *et al*., 2008). Comaniciu *et al* proposed the variable bandwidth mean shift by adaptively estimating a given histogram (Comaniciu, 2003). However, similar color may have different scales on different positions. Selection of kernel scale is important in the mean-shift algorithm, but there is no natural mechanism for choosing automatically the scale.

In this paper, we propose an algorithm for image segmentation using mean shift at multiple scales and apply to skin color image segmentation. The goal of the proposed algorithm is to segment images which include smooth but relatively fast changing illumination and to address its effects on the skin-color appearance.

ADAPTIVE MEAN SHIFT FOR IMAGE SEGMENTATION

Mean shift algorithm

The idea of mean-shift is to iteratively shift the data point to the average of the data points in a fixed-sized window (Comaniciu, 2003). It follows an estimation of the gradient of the density function without assuming any a priori distribution structure of the data. Let the probability density function $f(x)$ of a point x be unimodal. The multivariate bandwidth kernel density estimate is defined as (Silverman, 1986):

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

$$\overline{f}(x)=\frac{1}{nr^{d}}\sum_{i=1}^{n}K\left(\frac{x-x_i}{r}\right) \tag{1}$$

where $\{x_i\}_{i=1\cdots n}$ represent a random sample from some unknown density *f*. The kernel *K* is taken to be a radially symmetric, nonnegative function satisfying

$$K(x)=c\prod_{i=1}^{d}k\left(\|x\|^{2}\right)$$

where *c* is the normalization constant and *d* is the number of dimension. The kernel function is centered at zero and integrating to one. A fixed search window radius *r* is a constant across $x\in r^{d}$ and the density at each *d*-dimension point *x* is estimated by taking the average of identically scaled kernels centered at each of the data points.

A differentiable kernel can define the estimate of the density gradient as the gradient of the kernel density estimate, Equation (1),

$$\overline{\nabla} f(x)\equiv\nabla\overline{f}(x)=\frac{1}{nr^{d}}\sum_{i=1}^{n}\nabla K\left(\frac{x-x_i}{r}\right) \tag{2}$$

For the Epanechnikov kernel, the density gradient estimates, Equation (2), is computer as

$$\overline{\nabla} f(x)=\frac{n_x}{n\left(r^{d}c^{d}\right)}\frac{d+2}{r^{2}}\mu_r(x) \tag{3}$$

where c^{d} is the volume of the unit *d*-dimensional sphere, a data point n_x is in the search window $S_r(x)$, and $\mu_r(x)$ is called the *sample mean shift*.

Because of the quantity $\frac{n_x}{n\left(r^{d}c^{d}\right)}$ is the kernel density estimate $\overline{f}(x)$, from Equation (3), we can compute the sample mean shift as an offset to shift by

$$\mu_r(x)=\frac{r^{2}}{d+2}\frac{\overline{\nabla} f(x)}{\overline{f}(x)} \tag{4}$$

Note that an estimate of the normalized gradient can be obtained by computing the sample mean shift in a uniform kernel centered on *x*. Thus, the mean-shift vector, the vector of difference between the local mean and the center of the window, is proportional to the gradient of the probability density at *x*. This is a beneficial when the highest density region of the probability density function is sought. Such region corresponds to small mean shift. On the other hand, low-density regions correspond to large mean shifts. The shifts are always in the direction of the probability density maximum, the mode. At the mode the mean shift is close to zero. This property can be exploited in a simple, adaptive steepest ascent algorithm.

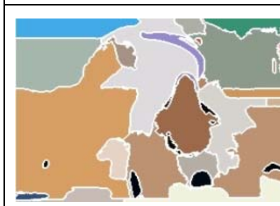
Algorithm I. Mean shift algorithm

1. Choose the radius *r* of the search window S_r .
2. Choose the initial location *x* in the window $S_r(x)$.
3. Compute the mean shift vector $\mu_r(x)$ and translate the search window $S_r(x)$ by that amount $\mu_r(x)$.
4. Repeat till convergence.

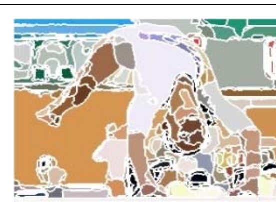
The mean shift always points to the direction of the maximal increase of the density; it can define a path leading to a local density maximum, *i.e.*, the mode of the density. The mean shift algorithm can move gradually towards the mode of a data set from any starting point (Algorithm I). Given a search radius and starting points, the mean shift algorithm can take place iteratively until convergence. The number of modes and position of each mode can be obtained simultaneously. Then the Euclidean distance of each sample is computed to all the detected modes and the sample is classified to the nearest.

Scale selection in mean shift

(a) original image



(b) under-segmentation



(c) over-segmentation

Figure-2. Image segmentation.

Scale is a crucial parameter to the performance of the mean shift algorithm and is directly determined by a

VOL. 10, NO. 16, SEPTEMBER 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

window radius r, (step 1 in Algorithm I). While results obtained appear satisfactory, when the local characteristics of the feature space differ significantly across data, it is difficult to find an optimal global scale for the mean shift algorithm. For image segmentation, if r is chosen too big, the too large window may contain many less-homogeneous pixels as well as homogeneous pixels. It causes convergence to a region between multiple models, rather than converging to just one of the modes, called under-segmentation. Figure-2(b) shows the left leg and left arm are clustered with the part of background and thus skin segmentation includes non-skin pixels. A too small sized window can roam around on a probability density plateau around the mode, called over-segmentation. Figure-2(c) shows all skin regions are segmented from the background, while most skin regions, especially the face and lower left leg, are divided into several parts as the mean shift window moves around the density plateau.

There is no natural mechanism for choosing or adapting r. Even when the best r is found, it is a global parameter. For good segmentation of natural images with varying illumination, variable radii are needed. The sample point density estimator, Eq. (1), can be defined by using of different radii for each data point x_i

$$\overline{f}(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{r(x_i)^d}K\left(\frac{x-x_i}{r(x_i)}\right) \tag{5}$$

which the estimate $\overline{f}(x)$ is the average of differently scaled kernels centered at each point.

Adaptive mean shift

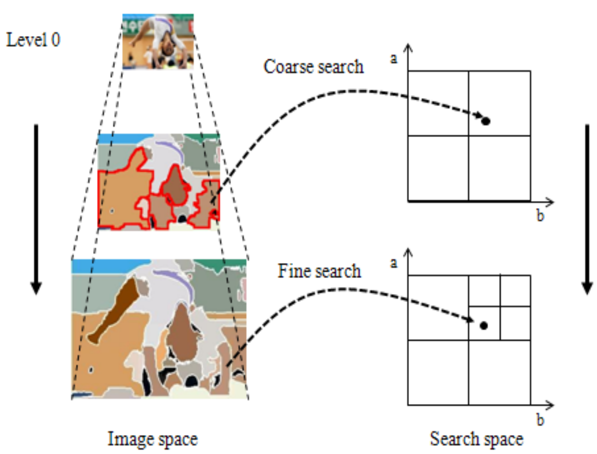


Figure-3. Adaptive mean shift using multi-resolution search for skin region segmentation: the original image is segmented with coarse search and then the result skin regions, drawn by thick lines, are refined with fine search.

To find such a variable scale, this paper suggests incorporating multi-resolution search. This use of a multi-resolution, coarse-to-fine search of density space to find appropriate plateau, is a general technique to prune the search space. Find search proceeds at only higher resolution. Figure-3 shows the outline of coarse- to-fine search to segment skin regions in various illuminations and skin color background. Coarse-to-fine search starts initially at lowest resolution (smallest size) first (step 1 in Algorithm II) and then proceeds to higher resolution where the search results are only refined (step 4 in Algorithm II). Generally, the higher (or coarser) resolution search results in under- segmentation. The details can be found down the image at higher resolution.

During the process of projecting the disparity map from the current level of the pyramid to the next (if the current level is not level 0), the image size was scaled up by the value of t (reduction ratio), and the disparity value by the same t. The disparity value, where the position of the new image is not a multiple of t, was obtained by linear interpolation.

Algorithm II. Adaptive mean shift algorithm.

1. Initialize the scale as the highest r, applied regions as a whole image, and the reduction ratio t.
2. Perform the mean shift algorithm with r to regions.
3. If all regions are well-segmented, stop.
4. Decrease r with the reduction ratio t; go to step 2.

Skin color image segmentation

In this section, we apply the proposed adaptive mean shift algorithm to segment a skin color image. At first, we smooth a RGB color space and remove noises in the image using a median filter. To make a statistical skin color model, histograms for the skin distributions are learned off-line from an image database (Jones and Rehg, 1999). Following Jones and Rehg, histogram-based skin distributions were computed at a 32×32×32 bin resolution in RGB color space. Conditional probability densities for skin color were obtained by dividing the count of pixels in each skin histogram bin *H(x)* by the total number of pixels in the skin histogram.

$$P(x \mid skin) = \frac{H(x)}{\sum_x H(x)} \tag{6}$$

Then, each pixel in an image is classified into skin color or non-skin color. Such a pixel-level classification has a major advantage, where the probability density function can be evaluated trivially regardless of the complexity of the underlying distribution. However, pixel-level classification is not enough because image background may also have skin colors.

VOL. 10, NO. 16, SEPTEMBER 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

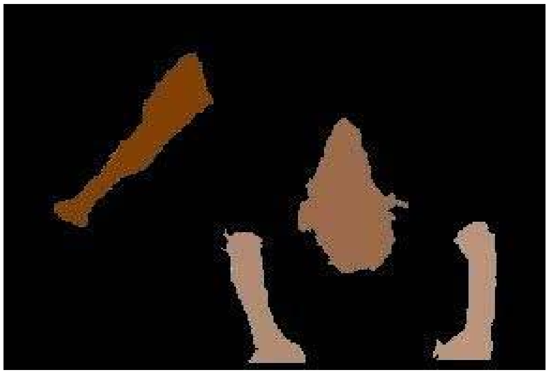


Figure-4. Skin region segmentation.

To resolve this problem, we use a region-based approach, mean shift clustering, which segments an image into regions and each region can compute a conditional probability density by averaging the density of each pixel, Eq. (6),

$$P(R \mid skin) = \frac{\sum_{x \in R} P(x \mid skin)}{Area(R)} \tag{7}$$

where *Aera*(*R*) is an area of a region *R* and is the total number of pixels in *R*. A skin region is defined as the area which has higher probability than a threshold value. Figure-4 presents the result of region-based approach to Figure-2(a).

EXPERIMENTAL RESULTS

The adaptive mean shift algorithm has been implemented and applied for skin region segmentation. In this paper, we collect 200 images which include various races, numbers of persons, and illuminations. For comparative experiments, we used the expectation-maximization (EM) algorithm (Dempster *et al*., 1977; Lee, 2004).

Figure-5 presents skin segmentation results using the EM algorithm with allocation (EAM) and the proposed algorithm. To evaluate the segmentation performance of the proposed algorithm, we compute the segmentation rate of skin regions which is the ratio of pixels that well segment. Table-1 shows the average segmentation rate on the experimental images and the proposed method has better segmentation accuracy than the EM or EAM algorithm.

Table-1. Average skin segmentation rate.

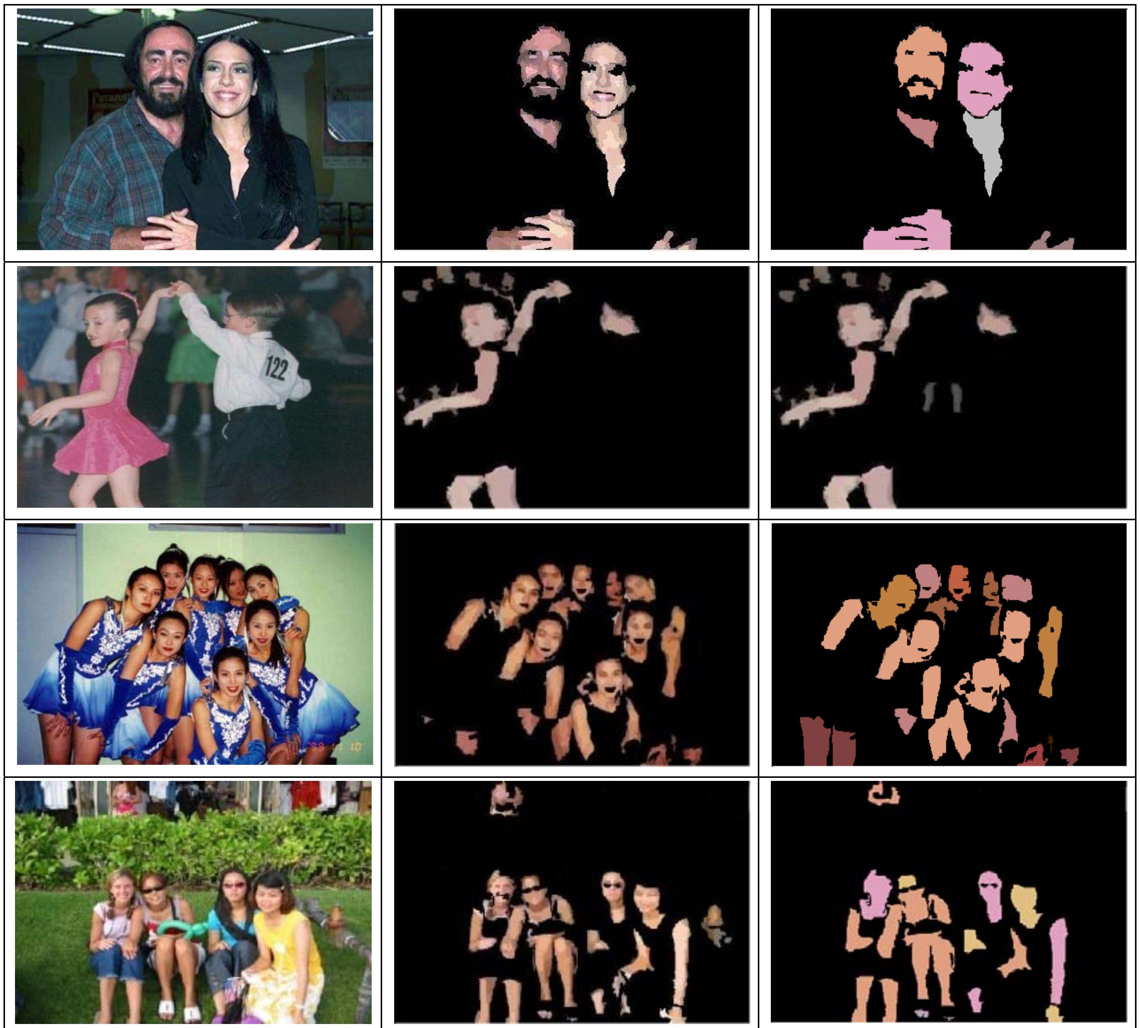
	Average skin segmentation rate
EM (Dempster *et al*. 1977) + Eq. (7)	86.15%
EAM (Lee, 2004) + Eq. (7)	92.35%
Our proposed method	94.75%

CONCLUSIONS

We have proposed an adaptive method that segments and applied to skin region segmentation. While the original mean shift algorithm uses a fixed size of the search window, a natural image contains various illuminations which need a various size of window. In this paper, we use multiple scales and coarse-to-fine search to find appropriate size of search window. We demonstrated that our coarse-to-fine approach has better segmentation.

Further direction of this work shall extend a pixel-based approach to region-based approach by including shape features. The proposed algorithm will be an important part of human detection and body posture recognition.

VOL. 10, NO. 16, SEPTEMBER 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences



(a) sample image	(b) EM with allocation (Lee, 2004)	(c) our proposed method

Figure-5. Experimental results.

REFERENCES

Pal N.R., Pal S. K. 1993. A review on image segmentation techniques. Pattern Recognition. 26(9): 1277-1294.

Comaniciu D, Meer P. 2002. Mean shift: A robust approach toward feature space analysis. IEEE transactions on Pattern Analysis and Machine Intelligence. 24(5): 603-619.

Mohammadi S. A., Mahzoun M. R. 2012. A novel approach coloured object tracker with adaptive model and bandwidth using mean shift algorithm. Signal and Image Processing: An International Journal. 3(3): 1-15.

Collins R. T. Mean-shift blob tracking through scale space. 2003. Proceedings of the IEEE international conference on computer vision and pattern recognition. pp. 234-240.

Fricker R. D. Jr., Knitt M. C., Hu C. X. 2008. Comparing directionally sensitive MCUSUM and MEWMA procedures with application to biosurveillance. Quality Engineering. 20(4): 478-494.

Comaniciu D. An algorithm for data-driven bandwidth selection. 2003. IEEE transactions on Pattern Analysis and Machine Intelligence. 25(2): 281-288.

Silverman B. W. 1986. Density estimation for statistics and data analysis. Chapman and Hall.

Jones M. J, Rehg J. M. 1999. Statistical color models with application to skin detection. Proceedings of the IEEE

ARPN Journal of Engineering and Applied Sciences

international conference on computer vision and pattern recognition. 274-280.

Dempster A. P., Laird N. M., Rubin D.B. 1997. Maximum-likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society - Series B. 39(1): 1-38.

Lee K. M. 2004. Elliptical clustering with incremental growth and its application to skin color region segmentation. Journal of Korean Information Science Society. 31(9): 1161-1170.