International Journal of Information Technology, Vol. 12 No. 2 2006

A Segmentation Approach for Natural Images

Yiping Hong[1], Jianqiang Yi[1], Dongbin Zhao[1], and Xinzheng Li[1]

[1] Laboratory of Complex Systems and Intelligence Science, Institute of Automation, Chinese Academy of Sciences, 100080, Beijing, China
{yiping.hong, jianqiang.yi, dongbin.zhao, xinzheng.li}@mail.ia.ac.cn

Abstract

This paper presents an improved segmentation approach derived from mean shift for natural images. The optimal color bandwidth under Plug-in rule is not always satisfying in the actual vision tasks, and a changing color bandwidth is helpful for controlling the segmentation result. The performance of direct density searching is better than mean shift under the same spatial bandwidth. A global optimization criterion for mode merging stabilizes the result in segmenting different images. Merging of texture regions mostly eliminates the influence from texture features. Based on the adjustable color bandwidth, direct density searching and a global criterion, the improved clustering approach performs better than mean shift, shown from experimental results. In addition, an image is partitioned into some local patches after mode detection. These patches can be taken as the initial segmentation for further processing that is based on a global optimization criterion with texture or statistical features.

1 Introduction

The existing segmentation approaches for natural images can be classified from different aspects, such as local vs. global, contour-based vs. region-based, applying to the intensity images vs. to the texture images, etc. The approaches based on local decision segment images using the pixel gradient or intensity similarity in a local window, while the ones based on global information often optimize a function constructed according to the global statistics. The contour-based and region-based approaches are not much different from each other since the boundaries of region can be defined as contour. The natural images often include rich texture, thus, the texture feature is necessary to combine with intensity information to get different semantic regions. The recent research work for natural image segmentation mainly focuses on the color-contour methods that are proved to be more efficient than ones only based on pixel intensity [1-4].

Segmentation by synthesizing contour and texture is often a global optimization process. J. Shi and J. Malik in their literature [5] first proposed a graph partitioning method that aims at extracting the global impression of an image. In [2], J. Malik furthers his work to introduce the contour and texture into graph theoretic framework of normalized cuts. The extensive works are done by the following researchers trying to synthesize the brightness and texture features from different perspectives [6-8]. The experimental results in these literatures show the performance of them is good. Compared with above methods that are not dependent on the models, some model-dependent approaches based on global optimization, such as region competition [9-10] and DDMCMC [11], have got some good results too. They use the models to depict different types of regions, and then classify the lattice with these constructed models efficiently.

Although above-mentioned methods are shown effective, they usually consume much computation to get the final results. The global optimization is an iterative process, which limits the application of it to some actual cases. Different from global optimization, mean shift clustering is a local approach, and it groups the pixels from the view of data analysis. The segmentation approach using mean shift clustering proposed by D. Comanicu in [12-13] only depends on the intensity of pixels without texture features, and groups the features with variable bandwidth. The experimental results in [13] show that the performance of mean shift is good although it has not considered the texture information. The more important is that mean shift clustering needs much less computation than global optimization approaches, and it can be used in real time video processing, such as object tracking [14], which makes it more applicable in some cases than global optimization approaches. However, on segmenting the texture images, it is not so effective as the global optimization approaches, and often over-segments the images.

To improve the performance of mean shift in segmentation of natural images with texture features, in this paper, a segmentation approach derived from mean shift is proposed. In this proposed approach, bandwidth selection, the way of searching local modes and the merging criterion of detected modes are different from mean shift. It considers more about segmentation of images with texture and control of segmentation result.

The outline of this paper is as follows. In Section 2 is the overview of mean shift clustering. Section 3 describes in detail the improved clustering approach including bandwidth selection, mode detection, mode merging and merging of texture regions. Section 4 shows some experimental results, and the conclusion is drawn in Section 5.

2 Overview of Mean Shift Clustering

Fukunaga first introduced mean shift algorithm into pattern recognition field [15]. The advantage of employing mean shift procedure in clustering was only recently rediscovered [16]. The literature [12] has done some further researches on bandwidth selection. A recent comprehensive research on applying mean shift to image segmentation refers to [13].

Mean shift always clusters data along the incremental direction of density. The relation between mean shift vector $M_{h,G}(X)$ and density gradient vector $\hat{\nabla} f_{h,K}(X)$ is [13]

$$M_{h,G}(X) = \frac{1}{2} h^2 c \frac{\hat{\nabla} f_{h,K}(X)}{\hat{f}_{h,G}(X)}, \tag{1}$$

where $\hat{f}_{h,G}(X)$ is the estimated density, h is the bandwidth, and c is a constant. It is shown that the direction of $M_{h,G}(X)$ and that of $\hat{\nabla} f_{h,K}(X)$ are consistent. Therefore, to get the direction of $\hat{\nabla} f_{h,K}(X)$, it only need to calculate $M_{h,G}(X)$. $M_{h,G}(X)$ equals

$$M_{h,G}(X) = \frac{\sum_{i=1}^{n} X_i g\left(\left\| h^{-1}(X - X_i)\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\| h^{-1}(X - X_i)\right\|^2\right)} - X, \tag{2}$$

where $k(u)$ is a kernel function, and $g(u) = -k'(u)$. $M_{h,G}(X)$ is the mean of the weighted vectors in the window centered on the point X.

Mean shift in segmentation should search the position of next trajectory point that has been quantized into integer according to $M_{h,G}(X)$ which is always a float vector. Meanwhile, a threshold to $M_{h,G}(X)$ should be set to stop shift. This leads to blur the regions with high density. The bandwidth h is calculated based on Plug-in rule by minimizing the Asymptotic Mean Integrated Square Error (AMISE), and it is global variable or fixed. Combing the detected local modes whose color is within the bandwidth can lead to the final segmentation result.

3 Improved Clustering Approach

In mean shift clustering, after mode detection, an image can be partitioned into a series of local modes. Assuming that an image is composed of K local modes, denoted by $\theta^{(k)}$. Then, detection of local modes at the arbitrary point X is

$$\hat{\theta}^{(k)} = \arg\max_{X}{}_k \; f(X \mid X_1, X_2, \cdots, X_n) \qquad k = 1, 2, \cdots, K, \tag{3}$$

where $f(X \mid X_1, X_2, \cdots, X_n)$ is a conditional density function.

3.1 Mode Detection Scheme

Mean shift clustering always shift along the direction of density increasing which is represented by $M_{h,G}(X)$, referring to the equ. (2). The shift procedure for mode detection is also the procedure of blurring the images and is similar to filtering. The reason for using $M_{h,G}(X)$ in shift procedure is that it consumes less computation than directly searching the direction of density increasing with the same search window. However, using the same window to search local modes, direct density searching performs better than mean shift clustering. Here, the way of direct density searching for mode detection with a small search window is adopted.

Let $X_i = (x_{i1,} \cdots, x_{id}), i = 1, \cdots, n$, and $\{X_1, X_2, \cdots, X_n\}$ is a d-dimensional sample set. Given a d-dimensional stochastic variable X , its nonparametric kernel density estimator is

$$\hat{f}_H(X) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{\det(H)}K\left\{H^{-1}(X - X_i)\right\} = \frac{1}{n}\sum_{i=1}^{n}K_H\left\{(X - X_i)\right\}, \tag{4}$$

where H is a bandwidth matrix, and $K_H(\cdot) = \frac{1}{\det(H)}K(H^{-1}\cdot)$. $K(\cdot)$ is a radially symmetric, non-negative, monotonic, d-variate kernel function, and satisfies $\int K(x)dx = 1$ and $K(u) = 0 \quad if \quad |u| > 1$ [17-19].

In density estimation of an image, n means the total number of pixels. Because of $K(u) = 0 \quad if \quad |u| > 1$, only those pixels within the bandwidth matrix H act on the equ. (4). At each pixel, an estimated density is obtained.

Let $\hat{f}_H(X_i)$ be the estimated density at the location $\vec{x} = (x_1, x_2)$, and $\hat{f}_H(X_i^1), \hat{f}_H(X_i^2), \cdots, \hat{f}_H(X_i^p)$ be the estimated densities in the domain where p is the number of pixels within H . Then, the local extreme value of density in the domain can be determined according to

$$\hat{f}_i^1 = \max_{j \in S_{X_i}} f_H(X_i^j) * K_d\{H^{-1}(X_i^j - X_i)\},$$
$$if \quad \hat{f}_H(X_i^j) > \hat{f}_H(X_i), j \in 1, 2, \cdots, p, \tag{5}$$

where S_{X_i} is the domain at $\vec{x}$. $K_d(H^{-1}\cdot)$ is a distance function which is different from the kernel function $K_H(\cdot)$ and satisfies $K_d(u) = 0 \quad if \quad |u| > 1$. Recursively Shifting from $\vec{x}$ to the pixel $\vec{x}_i^1$ whose estimated density is $\hat{f}_i^1$ and searching local extreme point of density at new location can get a series of local extreme value $\hat{f}_i^1, \hat{f}_i^2, \cdots, \hat{f}_i^q$ and a series of locations $\vec{x}_i^1, \vec{x}_i^2, \cdots, \vec{x}_i^q$.

Because the series of $\hat{f}_i^1, \hat{f}_i^2, \cdots, \hat{f}_i^q$ is incremental and the upper limit of density is 1, $\hat{f}_i^1, \hat{f}_i^2, \cdots, \hat{f}_i^q$ will finally converge to the location $\vec{x}_i^q$. All pixels converging to the same $\vec{x}_i^q$ form a local mode, denoted by $\theta^{(k)}, k = 1, 2, \cdots, K$ where K is the number of modes in an image. A $\theta^{(k)}$ is often composed of many series of locations $\vec{x}_i^1, \vec{x}_i^2, \cdots, \vec{x}_i^q$. The color value of pixels within the same $\theta^{(k)}$ is replaced with that of $\vec{x}_i^q$. Therefore, after detection of local modes, an image can be decomposed into many patches, each of which corresponds to a $\theta^{(k)}$.

A 5-dimensional bandwidth matrix is adopted in this paper. Let $H = diag(h_x, h_y, h_l, h_u, h_v)$, where $h_s = h_x = h_y$ is the spatial bandwidth that is equal to the diameter of circle searching window, and h_l, h_u, h_v are the color bandwidths of LUV components [12-13]. From above mode detection scheme, it is known that direct density searching for mode detection contains two steps: density estimation for each pixel with h_s and searching maximum of estimated densities with h_s at each pixel. Mean shift clustering directly calculates the mean shift vector based on the pixel color at each pixel with h_s. It is evident that direct density searching takes two steps to complete the searching procedure, while mean shift takes one step to get the searching result. Therefore, the direct searching way is more time-consumed than mean shift, although any step in direct way is simpler than mean shift. However, the direct way searches the local density maximum by using color information of pixels within $2h_s$ window, and mean shift only depends on pixels within h_s window. This means the direct way will get a better searching result than mean shift at the cost of increasing computation. Meanwhile, if the size of h_s in direct way decreases, its computation will rapidly be cut down too, while its performance may be not worse than mean shift.

In mode detection, the bandwidth H is also an important parameter. A big color bandwidth h_l, h_u, h_v will lead to ignore image details and make the number of detected local modes decreased. The bandwidth h_l, h_u, h_v can be viewed as the resolution of segmentation. The number of detected local modes is also sensitive to noise. Therefore, filtering is a necessary step before mode detection.

3.2 Bandwidth Selection

The most important parameter in mean shift is the bandwidth matrix H. Plug-in and Cross-validation are two presently proverbially used bandwidth selection methods. In actual applications, it is difficult to say which method is better than the other [17-19]. The idea of Plug-in method is replacing the unknown parameter with the estimated one. If the kernel function is a Gaussian multivariate function, the conducted optimal bandwidth by Plug-in is [17]

$$\hat{h}_j = \left(\frac{4}{d+2}\right)^{1/(d+4)} n^{-1/(d+4)} \sigma_j , \tag{6}$$

where σ_j is the standard deviation, $d = 5$ is the number of dimension, and n is the sum of data.

Based on the equ. (6), h_l, h_u and h_v are estimated respectively by L, U, and V component of an image. σ_j is replaced with the estimate $\hat{\sigma} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$.

The selection of spatial bandwidth h_s relates to the actual vision task. When h_s increases, the consumed computation will rapidly go up and the detail of an image will be ignored. The experiments prove that a big spatial bandwidth h_s does not always improve the results. Therefore, for reducing the computation, a small value of h_s is preferable.

The color bandwidth h_l, h_u, h_v in equ. (6) are obtained by minimizing the mean integrated squared error. It is clear that the aim of choosing bandwidth h_l, h_u, h_v under Plug-in rule is to reduce the distortion between the estimated pdf and the original pdf as much as possible. The h_l, h_u, h_v is estimated from the view of data analysis. However, in computer vision, images should be segmented into semantic regions instead of keeping every detail. Sometimes, it is expected that only the outline of an image be captured after segmentation, which makes a small bandwidth is nonsense and a big bandwidth may be useful. Therefore, a changing bandwidth $k_h * h_l, k_h * h_u, k_h * h_v$ is more feasible in real vision systems, where k_h is an adjustable parameter.

3.3 Density Estimation

In density estimation, if the value of exponential function can be calculated beforehand, it will save much computation time and makes the whole segmentation fast. Let $H = diag\{h_1, h_2, \cdots, h_d\}$, the equ. (4) becomes [17]

$$\hat{f}_H(X) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_1 \cdots h_d} K\left(\frac{x_1 - x_{i1}}{h_1}, \frac{x_2 - x_{i2}}{h_2}, \cdots, \frac{x_d - x_{id}}{h_d}\right). \tag{7}$$

Usually $K(u) = k(u_1) \times k(u_2) \times \cdots \times k(u_d)$ where $k(u)$ is an univariate kernel function. Then,

$$\hat{f}_H(X) = \frac{1}{n}\sum_{i=1}^{n}\left\{\prod_{j=1}^{d} h_j^{-1} k\left(\frac{x_j - x_{ij}}{h_j}\right)\right\}. \tag{8}$$

In (8), after the kernel function $k(u)$ and the bandwidth h_j are selected, the value of $\hat{f}_H(X)$ is determined only by $x_j - x_{ij}$. Meanwhile, the range of $|x_j - x_{ij}|$ is between 0 and 512 in LUV space. Therefore, to cut down the computation in density estimation, a series of densities $\hat{f}_H(X)$ with $|x_j - x_{ij}|$ increasing from 0 to 512 by interval 0.1 is calculated and organized in a table beforehand. Then, in actual density estimation, $\hat{f}_H(X)$ can be looked up by $|x_j - x_{ij}|$ from the table.

3.4 A Global Criterion for Mode Merging

Mean shift clustering gets final segmentation result by merging the detected modes whose color is within color bandwidth. When the bandwidth is adjusted by the actual vision tasks instead of calculated by the equ. (6), the mode merging in mean shift is not suitable any more. Thus, a global evaluation criterion for mode merging is recommended here. Although a global criterion will lead the merging to be an iterative procedure, merging based on the detected modes instead of pixels makes its computation rapid.

Let $\overline{X}^{(k)}$ be the mean of color vectors in the k th region, and $\overline{X}$ be the mean of color vectors in the whole image. Then, the divergence J is defined as [20]

$$J = \frac{\sum_{i=1}^{K_m} \sum_{j=1}^{N_i} \left\| X_j^{(i)} - \overline{X}^{(i)} \right\|}{\sum_{i=1}^{n} \left\| X_i - \overline{X} \right\|}, \tag{9}$$

where K_m is the number of regions for merging and N_i is the number of pixels in i th region.

In merging, K_m is a monotonically decreasing variable while J is a monotonically increasing variable. The rule of stopping merging is to minimize $K_m + J$ until it reaches the minimal value.

The initial value of $K_m + J$ is determined by the initial value of K_m, denoted by K_m^0. If K_m^0 is big, image will be over-segmented. The number of local modes in an image is often not proper to be selected for K_m^0.

Let $h_{l0} = \alpha_0 \times h_l$, $h_{u0} = \alpha_0 \times h_u$, and $h_{v0} = \alpha_0 \times h_v$. Using h_{l0}, h_{u0}, h_{v0} as three thresholds of LUV color components to merge the local modes, the number of merged regions is denoted as K_m^0. Here, α_0 can be empirically selected according to the actual vision task. After K_m^0 is determined, let α_0 increase by step 0.2, i.e. $\alpha_{t+1} = \alpha_t + 0.2, t = 0,1,2,\cdots$. When $K_m + J$ reaches the minimal value, α_t becomes the optimal value α^* for merging. Let $h_l^* = \alpha^* \times h_l$, $h_u^* = \alpha^* \times h_u$, and $h_v^* = \alpha^* \times h_v$. Based on h_l^*, h_u^*, h_v^*, the final segmentation result will be obtained.

Here, only the parameter α_0 affects the final merged result, and it is selected by experiments. When α_0 increases the number of merged regions decreases, and vice versa. Therefore, α_0 can proximately control the number of merged regions.

3.5 Merging of Texture Regions

If an image includes texture features, after mode merging, there may still have some small patches resulting from texture that make the segmentation result unreliable. These small patches should be removed or merge to get a final result.

For each detected local mode $\theta^{(k)}$, a value can be calculated according to the following equ., which is called isolated value here, denoted by I^k.

$$I^k = \frac{N_i^k * M^k}{N^k} \quad , \quad M^k = (M_l^k + M_u^k + M_v^k)/3 \quad ,$$

$$M_j^k = \begin{cases} 1 & if \quad |x_{ij}^k - x_j^k| > h_j \\ \dfrac{|x_{ij}^k - x_j^k|}{h_j} & else \end{cases}$$

$$i = 1,2,\cdots,m^k,\ j = l,u,v \qquad (10)$$

where m^k is the number of local modes connected to the mode $\theta^{(k)}$ in spatial space, N_i^k is the number of pixels in the i th local mode connected to $\theta^{(k)}$, $N^k = \sum_i N_i^k$, $x_{ij}^k, j = l,u,v$ is the color component of pixels in the i th connected local mode, and $x_j^k, j = l,u,v$ is the color component of pixels in $\theta^{(k)}$.

Then, the mean and the variance of I^k for all local modes is calculated, denoted by $\bar{I}$ and I_σ. Supposes that there are m_R regions after mode merging, denoted by $R^m, m = 1,2,\cdots,m_R$, and the isolated value for R^m is IR^m. Since the region R^m is often composed of some local modes after mode merging, IR^m is defined being equal to the minimal value of I^k in the local modes that is within the same R^m. Then, the merging threshold for R^m is as follows:

$$T_I = \begin{cases} \bar{I} - I_\sigma & if \quad (\bar{I} - I_\sigma) > \alpha_I \\ \bar{I} & else \end{cases} \qquad (11)$$

where T_I is the threshold for merging R^m, α_I is also a threshold parameter. Sometimes, the value $\bar{I} - I_\sigma$ may be very close to zero but $\bar{I}$ and I_σ is big. Thus, α_I is set to avoid this case, and in this paper $\alpha_I = 0.1$. If $|IR^i - IR^j| > T_I, i, j = 1,2,\cdots,m_R$, these two regions will be merged.

From the equ (11), it is shown that if the color of two connected regions in R^m is quite different from that of the regions connected to them, they will be merged together. Because IR^m indicates the color variance in the linked local patches under the resolution of segmentation $k_h * h_l, k_h * h_u, k_h * h_v$, generally a low value of IR^m corresponds to a smooth surface in image, and a large value of IR^m often corresponds to a texture surface. Merging the regions R^m with big IR^m will proximately eliminate the over-segmentation resulted from texture feature. However, the merging does not distinguish the type of texture, but take them as the same object surface.

The region merging here is not based on the actual extracted texture feature. An essential resolution to the natural images segmentation with texture is to combine color, contour and texture feature simultaneously. However, the region merging approach in this paper can mostly remove the influence from texture, with a better performance than mean shift and a much faster speed than the approaches considering color and texture simultaneously.

3.6 The whole procedure of segmentation

The segmentation procedure of improved clustering approach is summarized as follows:

1) Estimating three color bandwidths h_l, h_u, and h_v.
2) Filtering.
3) Density estimation at each pixel.
4) Detection of local mode.
5) Choosing α_0 and merging local modes.
6) Merging of texture regions.
7) Eliminating the regions smaller than N_m pixels (optional).

4 Experiments

In the improved clustering approach, there are three parameters need to be determined. They are h_s, k_h and α_0. h_s mainly influences the computation speed. k_h adjusts the color bandwidth and is determined according to the actual applications. A big k_h will overlook much detail of an image in segmentation. α_0 affects the final number of merged regions, which can be determined by experiments.

The images in the top row in Fig. 1 show the edges detected by EDISON (A mean shift clustering software compiled according to the literature [13]) with $(h_s, h_c, N_m) = (7,14,200)$, where h_c is the color bandwidth. The images in the second row correspond to the improved approach with $(h_s, k_h, \alpha_0, N_m) = (5,2,0.4,200)$, and the parameters in the third row are $(h_s, k_h, \alpha_0, N_m) = (7,2,0.4,200)$. The segmented images by improved approach are all under the bandwidth $2*h_l, 2*h_u, 2*h_v$, direct density searching way and a global criterion for mode merging.

From Fig.1, it can be seen that the segmented images resulted from the improved approach are better than that of mean shift clustering. In the top row images, there are always some unexpected little patches that are seldom existed in the other images in Fig.1. However, there are almost not any small patches in all images segmented by the improved approach, which indicates that merging of texture regions has behaved effectively. To segment different images, the performance of mean shift with the same parameter greatly fluctuates. Some images are over-segmented and some images are not enough segmented. On the contrary, the improved approach performances stably. This means the global criterion for mode merging can effectively stabilize the segmentation result. The segmentation results in the second row and the third row are not quite different from each other. Therefore, the value of h_s does not influence the results greatly.

Fig.2 shows two images segmented with different bandwidth $k_h = 1$ and $k_h = 2$, and the other parameters are $(h_s, \alpha_0, N_m) = (5, 0.2, 200)$. It is clear that the bandwidth (resolution of segmentation) greatly influences the segmentation results. When the bandwidth goes up, more details in images will be ignored. Fig.3 is the segmentation results under different α_0 with $(h_s, k_h, N_m) = (5, 2, 200)$. It can be seen that α_0 can control the number of final merged regions, while k_h can control the resolution of segmentation in the whole image.

Using a computer with CPU 2.0G to segment the images whose size is 288*384 in Fig.1, the time consumed by improved clustering approach is less than 1 seconds when $(h_s, k_h, \alpha_0, N_m) = (5, 2, 0.4, 200)$, and less than 2 seconds when $(h_s, k_h, \alpha_0, N_m) = (7, 2, 0.4, 200)$. The C++ code has not been optimized. This indicates that the improved clustering approach is not much time-consumed, while its performance is more stable and more effective than mean shift. Compared with the segmentation approaches based on global optimization such as normalized cuts and region competition [2,9-10], it is much faster. In addition, the result of the improved approach can be easily controlled by changing k_h and α_0. With the code optimized, such as the value of exponential function can be calculated beforehand and organized into a table in density estimation, the time spent will be greatly cut down.

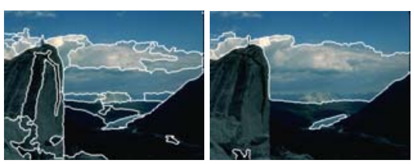
What is more, after processing with the improved clustering approach, an image is partitioned into some local patches. These patches can be used as the initial segmentation for an approach based on global optimization criterion with texture features.

5 Conclusion

Bandwidth selection, mode detection, a global criterion for mode merging and merging of texture regions are described in detail in this paper. The experimental results show that the segmentation performance has been improved. In the improved approach, the segmentation results can be easily controlled by changing color bandwidth. The direct density searching way does better than mean shift in segmentation, although it costs more computation under the same spatial bandwidth. A global criterion for mode merging is effective in stabilizing the results when different images are segmented. The merging of texture regions mostly eliminates the influence from texture.

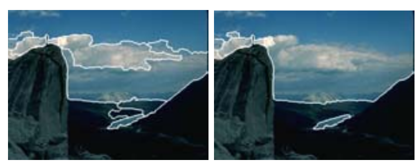
Three important parameters in the improved approach are h_s, k_h and α_0. h_s determines the segmentation speed. The choice of k_h relates to the actual vision task, and α_0 can be selected empirically. The experimental results indicate that $\alpha_0 = 0.4$ is a rational choice.

Fig.1. Comparison between mean shift and the improved clustering approach. The images in the first row are detected edges by EDISON (A mean shift clustering software). The images in the second row correspond to the improved approach with $(h_s, k_h, \alpha_0, N_m) = (5,2,0.4,100)$. The parameters for the images in the third row are $(h_s, k_h, \alpha_0, N_m) = (7,2,0.4,100)$.



(a) $k_h = 1$ (b) $k_h = 2$

Fig.2. Segmentation under different k_h



(a) $\alpha_0 = 0.2$ (b) $\alpha_0 = 0.6$

Fig.3. Segmentation under different α_0

References

1. D. Martin, C. Fowlkes, J. Malik: Learning to Detect Natural Image Boundaries Using Local Brightness, Color and Texture Cues. IEEE Transactions on PAMI. Vol. 26, no. 5 (2004) 530-549
2. J. Malik, S. Belongie, T. Leung, J. Shi: Contour and Texture Analysis for Image Segmentation. International Journal of Computer Vision. Vol. 43, no. 1 (2001) 7-27

3. J. Fan, D. Yau, A. Elmagarmid, W. Aref: Automatic Image Segmentation by Integrating Color-Edge Extraction and Seeded Region. IEEE Transactions on Image Processing. Vol. 10, no. 10 (2001) 1454-1466
4. Y. Deng, B. Manjunath: Unsupervised Segmentation of Color-Texture Regions in Images and Video. IEEE Transactions on PAMI. Vol. 23, no. 8 (2001) 800-810
5. J. Shi, J. Malik: Normalized Cuts and Image Segmentation. CVPR 1997. (1997) 731-737
6. K. Chen, S. Chen: Color Texture Segmentation Using Feature Distributions. Pattern Recognition Letters. Vol. 23, no. 7 (2002) 755-771
7. J. Chen, T. Pappas, A. Mojsilovic, B. Roqowitz: Image Segmentation by Spatially Adaptive Color and Texture Features. ICIP2003. Vol. 1 (2003) 1005-1008
8. J. Xu, P. Shi: Natural Color Image Segmentation. ICIP2003. Vol. 1 (2003) 973-976
9. S. Zhu, A. Yuille: Region Competition: Unifying Snakes, Region Growing, and Bayes/MDL for Multiband Image Segmentation. IEEE Transactions on PAMI. Vol. 18, no. 9 (1996) 884-900
10. M. Tang, S. Ma: General Scheme of Region Competition Based on Scale Space. IEEE Transactions on PAMI. Vol. 23, no. 12 (2001) 1366-1378
11. Z. Tu, S. Zhu: Image Segmentation by Data-Driven Markov Chain Monte Carlo. IEEE Transactions on PAMI. Vol. 24, no. 5, (2002) 657-673
12. D. Comaniciu, V. Ramesh, P. Meer: The Variable Bandwidth Mean Shift and Data-Driven Scale Selection. ICCV2001. Vol. 1 (2001) 438-445
13. D. Comaniciu, P. Meer: Mean Shift: A Robust Approach Toward Feature Space Analysis. IEEE Transactions on PAMI. Vol. 24, no. 5 (2002) 603-619
14. D. Comaniciu, V. Ramesh, P. Meer: Real-Time Tracking of Non-Rigid Objects Using Mean shift. CVPR 2000. Vol. 2 (2000) 142-149
15. K. Fukunaga, L.D. Hostetler: The estimation of the gradient of a density function, with applications in pattern recognition. IEEE Transactions on PAMI. Vol. 21 (1975) 32-40
16. Y. Cheng: Mean shift, mode seeking, and clustering. IEEE Transactions on PAMI. Vol. 17, no. 8 (1995) 790-799
17. W. Härdle, M. Müller, S. Sperlich, A. Werwatz: Nonparametric and Semiparametric Models. Springer, 2004
18. W. Härdle, L. Simar: Applied Multivariate Statistical Analysis. Springer, 2003
19. B.W. Silverman: Density Estimation for Statistics and Data Analysis. Chapman and Hall, London, 1986
20. Q. Ye, W. Gao, W. Wang, T. Huang: A color image segmentation algorithm by using color and spatial information. Chinese Journal of software. Vol.15, no. 4 (2004) 522-530

Yiping Hong received his Ph. D. degree in 2005 from Chinese Academy of Science, China, and currently is an image processing researcher in VIA Technologies, Inc. His main research interests include computer vision, image processing, robotics, etc.

Jianqiang Yi received the B.E. degree from Beijing Institute of Technology, China, in 1985, and the M.E. and Ph.D. degrees from Kyushu Institute of Technology, Japan, in 1989 and 1992. From 1992 to 1994, he joined Computer Software Development Company, Tokyo. From 1994 to 2001, he worked as a chief researcher at Mycom, Inc., Kyoto. Currently, he is a professor in the Laboratory of Complex Systems and Intelligence Science, Institute of Automation, Chinese Academy of Sciences. His research interests include theories and applications of fuzzy control, neural networks, intelligent control, intelligent robotics, and underactuated systems.

Donbing Zhao received his Ph.D. degree from Harbin Institute of Technology in 2000. He is now an associate professor in Institute of Automation, Chinese Academy of Sciences. His research interests include intelligent control, robotics, mechatronics.

Xinzheng Li received his Master degree in 2005 from Chinese Academy of Science, China. His main research interests include computer vision, image processing, robotics, etc.