

# 基于 RB-K 平均带宽设定的 Adaptive Mean shift

赵万磊 张学杰

(云南大学信息学院计算机科学系, 昆明 650091)

**摘要** Mean shift 是一个用在图像滤波、图像分割中的迭代过程。对于迭代过程, 带宽的设定很重要。固定带宽的 Mean shift 方法对于输入数据适应能力差, 而现有可变带宽估计方法有的对于固定带宽的方法改进并不明显, 有的时间复杂度较高, 因此对现有带宽估计方法做了改进, 提出了一种新的带宽估计方法, 即采用 K 平均聚类算法进行带宽估计, 并通过实验证实了该方法的有效性。

**关键词** Mean shift K 平均 图像分割 图像滤波

中图分类号: TP391 文献标识码: A 文章编号: 1006-8961(2006)04-0511-05

## RB-K-means Based Adaptive Mean shift

ZHAO Wan-lei, ZHANG Xue-jie

(Department of computer science, School of Information Science, Yunnan University, Kunming 650091)

**Abstract** Mean shift is an effective iterative procedure that has been used in edge preserving filtering and image segmentation. It is critical to set bandwidth properly for iteration. However, traditional fixed bandwidth mean shift is sensitive to input data, and current variable bandwidth estimation methods are not popularly accepted for their high time complexity or inconvenience. Here a new bandwidth estimate method has been proposed, namely, K-means is employed to estimate bandwidth for each sub-region and its effectiveness is indicated by experiments.

**Keywords** Mean shift, K-means, image segmentation, image-filtering

## 1 引言

Mean shift 方法最早由 Fukunaga 和 Hostetler 提出<sup>[1]</sup>, 它现在主要用于图像分割<sup>[2,3]</sup>、图像滤波<sup>[2,3]</sup>以及视频跟踪<sup>[4]</sup>, 并可以获得比较满意的效果。Mean shift 虽然实现了无参聚类或滤波, 但其迭代过程需要预先为每个需迭代的初始向量设定带宽 (Bandwidth)。现有设定带宽的方法是固定带宽法 (即为对所有的向量的迭代设定同一个带宽), 然而为所有向量找到一个合适的带宽却很困难; 而数据驱动 (Data Driven) 的可变带宽估计方法, 有的方法依赖于参数, 有的过于耗时<sup>[3]</sup>。为使 Mean shift 迭代过程可以根据不同的数据区域特征选择合理的带

宽, 同时兼顾算法的运行效率。本文采用 K 平均聚类算法分析数据集各个区域, 将数据集划分为一些区域, 然后根据各区域特征设定带宽。这一方法降低了现有可变带宽方法的时间复杂度, 而且使 Mean shift 方法的应用更具灵活性。

## 2 Mean shift 方法简介

### 2.1 Mean shift 方法原理

对于一个欧几里德空间内的点集, 无参密度估计方法根据一个点周围一个邻域内点的分布情况来估计空间中该点位置的密度; 类似地, 对空间中某一位置密度梯度的估计可以通过采样该位置的一个邻域来进行<sup>[1]</sup>。Fukunaga 和 Hostetler 根据这一思

基金项目: 云南省科技攻关项目 (2002GG15); 云南省省院省校合作项目 (2001JAALA02A002)

收稿日期: 2005-03-18; 改回日期: 2005-06-07

第一作者简介: 赵万磊 (1979 ~ ), 男, 现为云南大学信息学院计算机应用专业硕士研究生。主要研究方向为文本聚类、图像滤波、图像分割。E-mail: zhaowanl@yahoo.com.cn

路获得了 Mean shift 迭代过程。

设  $o$  是欧几里德空间中的一个点,  $C$  是以  $o$  为圆心,  $h$  为半径的圆形区域。Mean shift 根据空间中  $C$  区域内的点估计  $o$  处的密度梯度。然后, 迭代过程将  $o$  按该梯度的大小和方向移动至新位置  $o'$ , 接着根据  $o'$  为圆心的圆形邻域估计  $o'$  处的密度梯度, 再将  $o'$  按该梯度值的大小和方向移动, 如此迭代, 直到新位置处梯度估计值为 0。图 1<sup>[2]</sup> 演示了 2 维欧几里德空间中, 一个点采用 Mean shift 方法沿梯度方向在 2 维空间中的迭代移动的轨迹。图中的圆形区域是迭代起始点用于估计该点处密度梯度的邻域, 圆形区域的半径  $h$  称为带宽。

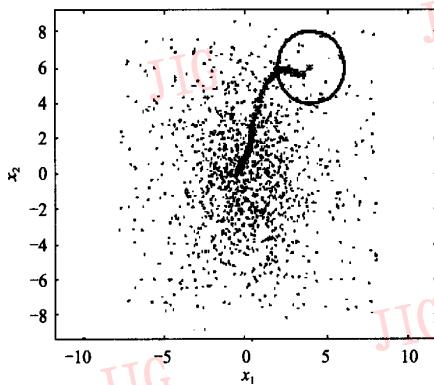


图 1 一个点的迭代轨迹

Fig. 1 Mean shift iterative trail of one point

若对空间中的每一点使用上述迭代过程, 那么散布在空间中的点将聚集到一些模式点 (mode) 处, 这些模式点处的密度梯度为 0。这样, 就可以将移动至同一个模式点处的点视为同类, 事实上就将空间中的点做了聚类<sup>[1]</sup>; 另外, 如果对一幅图像中的一个像素点使用 Mean shift 迭代过程, 然后用迭代后的模式点替代初始点, 就获得了这幅图像的滤波图。实验表明, 这是一种保持图像中物体边界 (Edge preserving) 的滤波, 而且可以用它进一步来做图像分割<sup>[2]</sup>。

下面给出 Mean shift 方法的严格定义。给定在  $d$  维的欧几里德空间  $\mathbf{R}^d$ , 对于数据集  $S = \{x_1, x_2, \dots, x_i, \dots, x_n\}$ 。点  $x$  处的密度估计函数<sup>[2]</sup>为

$$f(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \quad (1)$$

其中,  $h$  称为带宽, 它表明在多大的  $x$  邻域内估计  $x$  点处的密度;  $K(x)$  被称为密度核函数。对式 (1) 进行微分, 可以得到  $x$  处的密度梯度, 从而得到 Mean

shift 迭代向量<sup>[2-4]</sup>:

$$M_v(x) = \frac{\sum_{i=1}^n \frac{x_i}{h^{d+2}} \mathcal{G}\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n \frac{1}{h^{d+2}} \mathcal{G}\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} - x \quad (2)$$

其中,  $g(x) = -K'(x)$ 。式 (2) 的意义是, 如果要让  $x$  向带宽  $h$  范围内密度最大的地方移动, 则沿  $M_v(x)$  方向移动是最快的。由式 (2) 获得 Mean shift 迭代过程如下:

```
do {
  compute  $M_v(x)$ ;
   $x = x + M_v(x)$ ;
} while ( $M_v(x) \neq 0$ );
```

还需确认这样不断迭代的过程会最终收敛, 即  $M_v(x)$  最终为 0。这一迭代过程的收敛性已经获得了证明<sup>[1]</sup>。

## 2.2 Mean shift 带宽的选择

对于 Mean shift 迭代过程, 带宽  $h$  是一个重要的参数, 它决定对于  $x$  点处密度梯度值的估计将在多大的范围内进行。带宽虽然不作为一个用户输入参数, 但预先设定带宽对于迭代结果起着重要作用。现有的带宽方法有: 固定带宽法<sup>[2]</sup>、Balloon 估计法<sup>[4]</sup>、样本点估计法<sup>[4]</sup>。除固定带宽方法外, 其他方法都属于可变带宽方法, 采用可变带宽的 Mean shift 迭代过程被称为 Adaptive Mean shift。

### 2.2.1 固定带宽

虽然采用固定带宽法将 Mean shift 应用到图像滤波和图像分割中可以得到较好的效果。然而, 这一方法的缺陷是显然的, 例如对于图像滤波, 不同的输入图像, 色彩分布截然不同, 从而很难为不同的输入图像都设定一个合理的带宽。因此, 根据数据具体区域的特征来设置带宽更为合理。

### 2.2.2 可变带宽

针对固定带宽方法的不足, 可变带宽方法可依据数据区域的局部特征为不同的点设置不同的带宽。现在主要有 Balloon 估计法和样本点估计方法。从式 (1) 可以衍生出两种变化带宽的方法。

其一是, Balloon 方法将带宽的选择视作是一个关于 Mean shift 当前迭代点  $x$  的函数:  $h = h(x)$ 。于是得到了采用这种带宽迭代方法的密度估计函数:

$$\hat{f}_1(x) = \frac{1}{nh(x)^d} \sum_{i=1}^n K\left(\frac{x-x_i}{h(x)}\right) \quad (3)$$

式 (3) 与式 (1) 的不同之处在于式 (3) 中  $h$  是关于  $x$

的函数。为此需要重新证明以式(3)为密度估计函数的 Mean shift 迭代过程是否收敛。由于在对一个  $x$  的迭代中  $h$  并不发生变化,因此它的迭代收敛性采用类固定带宽迭代收敛性证明方法就可以得到证明<sup>[2]</sup>。对于  $h = h(x)$  的选择,最简洁的方法是取  $x$  到第  $L$  个近邻的距离作为带宽。但是,因为对每一个  $x$  (一般是多维向量)的最近邻查询,即使采用现在最好的多维索引结构  $R$ -tree<sup>[5]</sup>、 $R^*$ -tree<sup>[6]</sup>、 $X$ -tree<sup>[7]</sup> 或者近似最近邻索引 LSH (local sensitive hashing)<sup>[3]</sup> 也要  $\log(n)$  的时间复杂度;对于  $n$  个点,求带宽的时间复杂度将为  $n\log(n)$ ,而且这样引入了近邻个数  $L$  这个参数。

其二是,将带宽看作是一个随  $x_i$  变化的量,即  $h = h(x_i)$ ,称之为样本估计法。其密度估计函数可表达为

$$\hat{f}_2(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h(x_i)^d} K\left(\frac{x - x_i}{h(x_i)}\right) \quad (4)$$

以该函数作为 Mean shift 核函数的迭代收敛性已获得证明<sup>[4]</sup>。这种方法的好处在于带宽本身反映了区域的密度分布,而且适当选择  $h(x_i)$  可以显著减少密度估计函数  $\hat{f}_2(x_i)$  与真实密度函数  $f(x)$  之间的偏差。一种合适的  $h(x_i)$  函数是<sup>[4]</sup>:

$$h(x_i) = h_0 \left[ \frac{\lambda}{f(x_i)} \right]^{1/2} \quad (5)$$

由于无法知道  $f(x_i)$ ,又需要先给出一个合适的带宽  $h_0$ ,采用式(1)获得一个密度估计函数  $\tilde{f}(x_i)$ ,替换式(5)中的  $f(x_i)$ 。然而正如前面所说,要找到一个合适的独立于输入数据的  $h_0$  很困难。因此这种方法仍然面临参数值难以选择的问题。

### 3 基于RB-K平均聚类的带宽设定方法

以上评价了两种可变带宽的设置方法。确定带宽的实质是确定哪些点纳入到迭代过程,来决定  $x$  点一次迭代中移动的步长和方向。因此,直观的想法是  $x$  邻域内且最终同  $x$  一样移动到同一模式点的那些点应该为带宽所包括,即这些点是同属于一类的点。可取类内距离(最大、最小、或平均距离)来近似带宽来囊括它们。需要一种聚类方法首先对数据集进行聚类分析,得到关于数据集的一个分类,从而进一步确定带宽。最近,K平均聚类算法有了新的改进,提出了 Repeated Bisecting K 平均聚类算法<sup>[8,9]</sup>。大量实验<sup>[8-10]</sup>表明,它具有以下很好的性

质:线性时间复杂度;良好的聚类结果;适应多种数据类型;对数据输入顺序不敏感。而且它使用了优化函数来衡量聚类结果的优劣也提高了聚类结果的稳定性。

下面将进一步说明采用这种带宽设置方法 Mean shift 迭代过程的收敛性。设已用 RB-K 平均方法对数据集聚类,对于一个簇,取簇内的平均距离为该簇中每个点的带宽。于是对于数据集  $S = \{x_1, x_2, \dots, x_i, \dots, x_n\}$ ,有带宽集  $h(x_i)_{i=1, \dots, n}$ ,这样事实上将带宽  $h$  看作一个关于迭代点  $x$  的函数。由于该方法的收敛性已获得证明,可以断定这样的带宽设置方法是可行的。

另外, RB-K 平均被应用于遥感卫星图像的分割,这种分割并不能利用图像的区域连通性<sup>[11]</sup>,而只是利用 RB-K 平均良好的聚类性能来识别图片中人肉眼难以辨别的区域。图2是一幅直接用 K 平均聚类方法的图像分割结果。可以看出这种分割还不能称之为真正意义上的图形分割。然而,聚类结果将一幅图像识别为零碎的多个区域却为图像中每个点的带宽设置提供了方便。

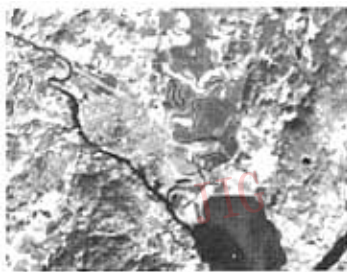


图2 K平均分割的遥感卫星图

Fig.2 Image segmented by K-means

### 4 基于RB-K平均的 Adaptive Mean shift 算法

#### 4.1 选择核函数及确定迭代过程

带宽设定后,还必须为 Mean shift 迭代过程选定合适的核函数。为了使 Mean shift 过程收敛,核函数的选择有一定要求,即设核函数为  $K(x)$ ,令  $K(x) = k(\|x\|^2)$ ,称  $k(x)$  为  $K(x)$  的轮廓函数。 $k(x)$  必须满足:非负性;非增性;分段连续且  $\int_0^{\infty} k(r) dr < \infty$ 。目前通常采用的核函数是 Gauss 核函数和 Epanechnikov 核函数<sup>[4,8]</sup>。这里采用比较简

洁的 Epanechnikov 核函数:

$$K_E(x) = \begin{cases} \frac{1}{2} c^{-1} (d+2) (1 - X^T X) & X^T X < 1 \\ 0 & \text{其他} \end{cases} \quad (6)$$

由于  $x_i - x \geq h$  时,核函数取值为 0。于是 Mean shift 向量可以进一步表示为<sup>[2]</sup>:

$$M_v(x) \equiv \frac{1}{n_x} \sum_{x_i \in S_{h_i}(x)} (x_i - x) = \frac{1}{n_x} \sum_{x_i \in S_{h_i}(x)} x_i - x \quad (7)$$

其中,  $S_{h_i}(x)$  是那些到  $x$  点距离  $\text{dist}(x, x_i) < h(x)$  的点集。 $n_x$  是  $S_{h_i}(x)$  中点的数目。

根据式(7),对于任意点  $x_j (j = 1, \dots, n)$ , Mean shift 的迭代过程 (MeanShiftIter) 变为

$$(1) I = 1; y_I = x_j;$$

$$(2) y_{I+1} = \sum_{x_i \in S_{h_i}(y_I)} x_i, I = I + 1$$

其中,  $x_i$  满足:  $\text{dist}(y_I, x_i) < h(y_I)$ ;

$$(3) \text{重复步骤 1, 2 直到 } y_{I+1} = y_I.$$

#### 4.2 算法

带宽设置方法确定之后,就获得了整个迭代算法。算法首先采用 RB-K 平均聚类算法获得对图像点集的一个分类。由于 K 平均方法难以利用空间信息,因此本文算法进一步将属于一个类的点集分割为空间区域上连通的区域,然后取区域内点的平均距离为属于该区域点的带宽,采用上节所述迭代过程对点集中的每一个点进行迭代。

输入:  $S$ —数据集,  $K$ —聚类数目

输出: 迭代后的  $S'$  集

(1)  $\text{ArrayList clusters} = \text{Rbkmean}(S, M)$ ;

(2) for each cluster in clusters

{  $\text{regions} = \text{regionDetector}(\text{cluster})$ ;

for each region in regions

{  $\text{bandwidth} = \text{getBandwidth}(\text{region})$ ;

for each Pointx in region

$\text{shiftedV} = \text{meanshiftIter}(\text{Pointx}, \text{bandwidth}, \text{window})$ ;

}

}

其中, Rbkmean 为 Rb-K 平均函数,用于对图像进行分割;函数  $\text{regionDetector}(\text{cluster})$  将聚类中一个簇分割为多个空间上连通的区域;面对同类且空间上连通的数据集  $\text{region}$ ,带宽可有多种取法:可以取区域内的平均距离,最大、最小距离等,在实验中取平均距离的效果最佳,这里用  $\text{getBandwidth}(\text{region})$  函数来表示;  $\text{Window}$  是当前点  $\text{Pointx}$  的空间邻域,如  $3 \times 3$  邻域,在我们的实验过程中  $\text{Window}$  始终为 6,

它是一个经验系数。

#### 4.3 算法时间复杂度分析

在上述算法中,第 1 步的时间复杂度为  $O(\ln)$ ,  $I$  为迭代次数;第 2 步事实上是对  $S$  中每个点都使用一次 MeanShift 迭代过程,迭代过程  $\text{MeanShiftIter}$  本身通常在几次内就收敛,过程中的  $\text{regionDetector}(\text{cluster})$  也是线性时间复杂度,因此整个算法的时间复杂度为  $O(\ln)$ 。这与一般 Mean shift 算法的时间复杂度  $O(n^2)$ <sup>[12]</sup>,及用最近邻近似带宽的方法(时间复杂度为  $O(n \log n)$ )要好得多。

#### 4.4 算法在图像处理中的应用

为了验证这种 Mean shift 方法的有效性,将其应用到图像滤波和图像分割中。在对图像的处理过程中,采用图像的  $L^*u^*v$  空间(对其他色彩空间也做了实验,如  $RGB$ ,  $HIS$ ,  $HSV$ ,  $I1I2I3$  结果都没有  $L^*u^*v$  空间好),因为在这个空间中 3 个分量彼此独立<sup>[11]</sup>。

采用 Mean shift 做图像滤波:对于输入的每一个点的  $L^*u^*v$  向量,采用上节算法进行迭代,从而获得一个新的  $L^*u^*v$  向量;用该向量替换图像该坐标点的原向量,就获得了滤波图像。

为表述方便,将本文提出的可变带宽滤波方法记为 RBK-B,  $k$  为聚类数目;固定带宽滤波方法记为 FIX-B,其中,  $\sigma_s, \sigma_c$  分别是 Mean shift 迭代过程在坐标空间  $x * y$  和色彩空间  $L^*u^*v$  的带宽。

图版 I 图 1(c)、图 2(c)、图 3(c) 是对图版 I 图 1(a)、图 2(a)、图 3(a) 分别采用 4.2 节所述算法迭代 6 次的结果,图版 I 图 1(b)、图 2(b)、图 3(b) 采用文献[13]得到。由图可见这样的过滤可以很好地保持不同物体区域的边界,同时又去除图像的纹理细节,这是一种保持边界过滤。采用固定带宽的方法也可以得到这样的结果,但达到同样的效果可能要许多次设置带宽,而且比起聚类数目带宽参数设置并不是一个直观参数。

当然有了上述的过滤结果很容易用它进一步做图像分割<sup>[1-3]</sup>。这里还可以采用另外的分割思路,由于单独使用 K 平均做图像分割并不理想(如图版 I 图 3(d)所示),这是因为 K 平均算法难以利用空间联通性<sup>[11]</sup>,但可以先用 Mean shift 迭代得到滤波(图版 I 图 3(c)),再用 K 平均做分割,便可得到良好的分割效果(图版 I 图 3(e))。这样既吸取了 K 平均做分割速度快、无需先验知识的优点;又通过采用 Mean 迭代过程间接地利用了图像的区域连通性。

## 5 结 论

对已有的 Mean shift 迭代过程中带宽的选择做了改进。把 RB-K 平均聚类算法应用到带宽选择中,得到了一种新的可变带宽迭代方法,这种方法时间复杂度低,而且对于图像处理来说,不但可以用它做图像滤波,而且可以进一步方便地采用 K 平均聚类方法做分割。实验显示,对于简单彩色图像,这种分割结果相当得好,而且与其他 Mean shift 过程做分割不同的是,本文的过程采用聚类数目做参数,这一参数设置比选择带宽更直观。

由于 K 平均做图像分割没有利用任何先验知识,因此对于复杂图像的分割仍然有困难,在将来的研究中,将寻找新的方法以改善 K 平均对上述滤波图分割的效果。

### 参考文献 (References)

- Keinosuke Fukunaga, Larry D Hosteler. The estimation of the gradient of a density function with applications in pattern recognition [J]. IEEE Transactions on information theory, 1975, 21 (1): 32 ~ 40.
- Dorin Comaniciu, Peter Meer. Mean shift analysis and applications [A]. In: Proceedings of IEEE International Conference Computer Vision[C], Kerkyra, Greece, 1999:1197 ~ 1203.
- Bogdan Georgescu, Ilan Shimshoni, Peter Meer. Mean shift based clustering in high dimensions a texture classification example[A]. In: Proceedings of Ninth IEEE International Conference on Computer Vision[C], Nice, France, 2003,1:13 ~ 16.
- Dorin Comaniciu, Visvanathan Ramesh, Peter Meer. The variable bandwidth mean shift and data-driven scale selection [A]. In: Proceedings of International Conference on Computer Vision (ICCV'01)[C], Vancouver, BC, Canada, 2001,1:438 ~ 445.
- Guttman A. R-Trees: A dynamic index structure for spatial searching [A]. In: Proceedings of ACM SIGMOD[C], Boston, USA, 1984: 47 ~ 57.
- Beckmann N, Kriegel H-P, Schneider R, et al. The R\*-tree: an efficient and robust access method for points and rectangles[A]. In: Proceedings of ACM SIGMOD Int'l. Conference on Management of Data[C], Atlantic City, New Jersey, US, 1990: 322 ~ 331.
- Stefan Berchtold, Daniel A Keim, Hans-Peter Kriegel. The X-tree: an index structure for high-dimensional data[A]. In: Proceedings of the 22nd VLDB Conference[C], Mumbai, India, 1996: 28 ~ 39.
- Zhao Ying, George Karypis. Criterion Functions for Document Clustering Experiments and Analysis[R]. Minneapolis, Minnesota, USA: Department of Computer Science & Eng University of Minnesota, 2001: 1 ~ 40.
- Zhao Ying, George Karypis. Evaluation of hierarchical clustering algorithms for document dataSet[A]. In: Proceedings of the Eleventh International Conference on Information and Knowledge Management [C], McLean, Virginia, USA: ACM Press, 2002: 515 ~ 524.
- Michael Steinbach, George Karypis, Vipin Kumar. A Comparison of Document Clustering Techniques [R]. Minneapolis, Minnesota, USA: Department of Computer Science & Eng University of Minnesota, 2000: 8 ~ 20.
- Cheng H D, Jiang X H, Sun Y, et al. Color image segmentation: advances and prospects [J]. Pattern Recognition, 2001, 34 (12): 2259 ~ 2281.
- Cheng Yi-zong. Mean shift, mode seeking, and clustering[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 17(8): 790 ~ 799.
- Chirs M. Christoudias, Bogdan Georgescu. EDISON [CP/OL], <http://www.caip.rutgers.edu/riul/research/code/EDISON/>, 2005-11-20.



(a) 原图 192×128



(b) FIX-B ( $\sigma_x, \sigma_y$ )=(12,8)



(c) RBK-B K=7

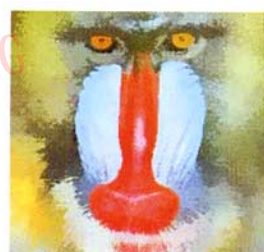
图1 过滤对比  
Fig.1 Comparison of filtering



(a) 原图 220×219

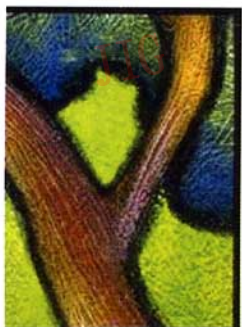


(b) FIX-B ( $\sigma_x, \sigma_y$ )=(12,8)



(c) RBK-B K=13

图2 过滤对比  
Fig.2 Comparison of filtering



(a) 原图 128×192



(b) FIX-B ( $\sigma_x, \sigma_y$ )=(12,8)



(c) RBK-B K=10



(d) 用RB-K平均分割(K=5)



(e) 用RB-K平均分割滤波(K=5)

图3 过滤及分割  
Fig.3 Filtering and segmentation