均值移动算法收敛性及均值移动矢量夹角分析

郭庆昌,王敏毅
GUO Qing-chang,WANG Min-yi

中国船舶重工集团公司 第七一〇研究所,湖北 宜昌 443000
CSIC-710 Research and Department Institute,Yichang,Hubei 443000,China
E-mail:boy_guoguo1@126.com

GUO Qing-chang,WANG Min-yi.Analysis on convergence of mean-shift and angle of continuous mean-shift vector. Computer Engineering and Applications,2010,46(7):34-38.

Abstract: Mean shift is an effective iterative algorithm.It has been widely used in the image processing and computer vision,but whose properties are not been perfectly proved.The convergence of the mean-shift is proved according to cauchy convergence criterion for the wrong and deficiency at the literatures.The angle of the mean shift vectors based on the arbitrary kernel in the same window is less than 90 degree.
Key words: mean-shift;convergence;Cauchy convergence criterion

摘 要:均值移动算法是一种统计迭代算法,目前在图像处理中得到了广泛应用。但是对其性质的分析仍然不完善,针对以往文献对均值移动算法收敛性证明的错误和不足,根据柯西收敛定理严格证明了均值移动算法的收敛性;证明了基于任意核,两连续均值移动矢量的夹角都不大于90°。
关键词:均值移动;收敛性;柯西收敛定理
DOI:10.3778/j.issn.1002-8331.2010.07.011 **文章编号:**1002-8331(2010)07-0034-05 **文献标识码:**A **中图分类号:**TP391.4

1975年Fukunaga和Hostetler在研究核函数的密度梯度估计时提出了均值移动算法的原型“valley-seeking procedure”[1]尽管效率非常出色,但并未得到学术界的注意。直到1995年Y.Cheng[2]成功将此算法扩展至计算机视觉领域,才引起了广泛的关注。开始Fukunaga等[1]认为均值移动算法是一种求最大值的最速上升法。1995年Y.Cheng[2]修正了Fukunaga的观点认为均值移动算法是具有自适应步长的最速上升法。2003年Y.Cheng等[3]提出了基于拟牛顿法的改进均值移动算法,使之达到了超线性收敛速度,他们的方法是直接计算梯度,用BFGS方法估计Hessian矩阵,计算迭代步长。2000年D.Comaniciu[4]等证明了均值移动算法密度随均值移动点单调递增性。2005年M.Fashing[5]等证明了当核函数为常数时,均值移动算法为边界优化算法。如果要对其进行改进,则需要收缩边界,找到合适的边界将耗费较多的计算量,但他们并没有给出改进的方法和途径。2005年李乡儒[6]基于多重假设证明了均值移动算法的收敛性,但是证明过程比较抽象使人不易理解。

近年来均值移动方法在计算机视觉和图像处理领域得到了广泛的应用,例如D.Comaniciu等[6-9]讨论了基于均值移动的目标跟踪算法;分析了均值移动算法的运动特征及其在图像分割中的应用[10-11];应用均值移动算法进行图像平滑[12];讨论了变窗宽均值移动算法[13-14];讨论了图像聚类中鞍点的解决方法[15]。G.R.Bradski等用均值移动算法进行人脸跟踪[16]。R.T.Collins等提出了变尺度目标跟踪算法[17]。Yang C等提出了基于核的目标跟踪算法[18-19]。朱胜利等用卡尔曼滤波器和均值移动算法相结合进行运动目标跟踪[20]。贾静平,张艳宁等提出了多自由度的均值移动目标跟踪算法[21]。李培华提出了一种改进的均值移动跟踪算法[22]。程伟、杨杰[23]用均值移动算法对红外目标进行跟踪。K.Deguchi[24]等采用粒子滤波器和均值移动算法相结合进行目标跟踪。A.Bandera等通过霍夫变换和均值移动算法相结合进行图像分割[25]等。Wang Yuzhong等[26]采用自适应均值移动算法和软规则对质地图像进行非监督分割。Wang J[27]等采用各项异性核进行图像分割和运动目标跟踪。M.Á等[28]采用高斯模糊均值移动算法进行非参数聚类。B.Georgescu[29]等在高维空间用均值移动算法聚类分析。H.Tek等[30]等基于射线传播规律采用均值移动算法进行血管检测。刘蓉等[31]用均值移动算法去除光谱信号中的噪声。

1 均值移动算法原理介绍

Mean-shift算法是一种基于核密度估计的无参快速模式匹配算法。利用核函数的性质,无需对整个区域的概率密度$\hat{f}_{h,k}(y)$进行估计,就能利用核对点的梯度进行估计,并进一步导出均值迁移步长。$K(x)$中心对称且满足$K(x)=c_{k,d}k(\|x\|^2)$,$k(x)$定义在$x\geqslant 0$区间上。$k(x)$称为$K(x)$的轮廓(profile)函数。于是核密度估计可以写成:

作者简介:郭庆昌(1979-),男,博士,工程师,主研领域为模式识别;王敏毅,男,教授,主研领域为无源光电干扰。
收稿日期:2008-09-23 **修回日期:**2008-12-22

$$\hat{f}_{h,k}(x)=\frac{c_{k,d}}{nh^d}\sum_{i=1}^{n}k(\|h^{-1}(x-x_i)\|^2) \tag{1}$$

使用核 $K(x)$ 的密度估计梯度为：

$$\nabla\hat{f}_{h,k}(x)=\frac{c_{k,d}}{nh^d}\sum_{i=1}^{n}\nabla k(\|h^{-1}(x-x_i)\|^2)=$$

$$\frac{2c_{k,d}}{nh^d}h^{-2}\sum_{i=1}^{n}(x-x_i)k'(\|h^{-1}(x-x_i)\|^2) \tag{2}$$

设 $g(x)=-k'(x)$，于是 g 对应一新核 $G(x)=c_{k,g}g(\|x\|)^2$。新核的概率密度为：

$$\hat{f}_{h,g}(x)=\frac{c_{k,g}}{nh^d}\sum_{i=1}^{n}g(\|h^{-1}(x-x_i)\|^2) \tag{3}$$

将 $g(x)$ 代入式(2)得：

$$\nabla f_{h,k}(x)=\frac{2c_{k,d}}{nh^d}h^{-2}\sum_{i=1}^{n}(x_i-x)g(\|h^{-1}(x-x_i)\|^2)=$$

$$\frac{2c_{k,d}}{nh^d}h^{-2}\left(\sum_{i=1}^{n}g(\|h^{-1}(x-x_i)\|^2)\right)\left(\frac{\sum_{i=1}^{n}x_ig(\|h^{-1}(x-x_i)\|^2)}{\sum_{i=1}^{n}g(\|h^{-1}(x-x_i)\|^2)}-x\right)$$

设 $$m_{h,g}=\frac{\sum_{i=1}^{n}x_ig(\|h^{-1}(x-x_i)\|^2)}{\sum_{i=1}^{n}g(\|h^{-1}(x-x_i)\|^2)}-x \tag{4}$$

式(4)为均值移动的移动矢量，结合式(3)得：

$$\nabla\hat{f}_{h,k}(x)=\frac{2c_{k,d}}{c_g}h^{-2}\hat{f}_{h,g}(x)m_{h,g}(x) \tag{5}$$

于是 $$m_{h,g}(x)=\frac{c_gh^2\nabla\hat{f}_{h,k}(x)}{2c_{k,d}\hat{f}_{h,g}(x)} \tag{6}$$

从式(6)可以看出 $m_{h,g}(x)$ 的运动方向与梯度的变化方向相同，大小与梯度的归一化值成正比。

均值移动算法的基于均值移动点的密度递增型在以往的文献[4]中已经进行了证明，这里不再说明。

2 均值移动算法收敛性

均值移动算法的收敛性就是证明均值移动函数的收敛性和均值移动矢量的收敛性。

均值移动算法是一种有效的统计迭代算法，而收敛性是任何迭代算法的必要前提。

以往文献对均值移动算法收敛性的证明分为以下3种：

（1）D.Comaniciu[32]等用条件"$\|y_{j+1}-y_j\|$ 收敛于0，则 y_j 收敛"证明收敛性；

（2）Meer[33]等用公式"$\|y_{j+m}-y_{j+m-1}\|^2+\cdots+$"$\|y_{j+1}-y_j\|^2\geqslant\|y_{j+m}-y_j\|^2$ 证明收敛性；

（3）李乡儒[6]根据多重假设证明 y_j 是收敛序列。

前两种情况的证明存在理论上的问题，文献[6]的证明过于抽象，使人很难理解。

对第一种情况：

例如：$y_j=\sum_{i=1}^{n}\frac{1}{i}$，则 $\|y_{i+1}-y_i\|=\frac{1}{i+1}$，当 $i\to\infty$ 时，$\|y_{i+1}-y_i\|\to0$，但是 y_j 不收敛。

对第二种情况：

例如：$y_3=3$，$y_2=2$，$y_1=1$

$\|y_3-y_2\|^2+\|y_2-y_1\|^2=2$

$\|y_3-y_1\|^2=4$

$\|y_3-y_2\|^2+\|y_2-y_1\|^2<\|y_3-y_1\|^2$ 可知命题错误。

例如：$y_3=[3,1,1,\cdots,1]$，$y_2=[2,1,1,\cdots,1]$，$y_1=[1,1,1,\cdots,1]$，每个都为 n 维向量。

$\|y_3-y_2\|^2+\|y_2-y_1\|^2=2$ $\|y_3-y_1\|^2=4$

$\|y_3-y_2\|^2+\|y_2-y_1\|^2<\|y_3-y_1\|^2$ 可知命题错误。

对第三种情况：

文献[6]的证明过程虽然正确，但是使人不易理解。

本节用柯西收敛定理[34]重新证明均值移动算法的收敛性，使人更容易理解。

柯西收敛定理 级数 $\sum_{n=1}^{\infty}u_n$ 收敛的充分必要条件为：对于任意给定的正数 ε，总存在自然数 N_z，当 $n>N_z$ 时，对于任意的自然数 p_z，都有 $|u_{n+1}+u_{n+2}+\cdots+u_{n+p_z}|<\varepsilon$ 成立。

定理1 当轮廓函数 $k(x)$ 为单调递减的凸函数时，密度函数序列 $\hat{f}_{h,k}(y_t)$ 和均值移动点列 $\{y_t\}_{t=0,1,\cdots,N}$ 收敛。

证明：

（1）证明密度函数 $\{\hat{f}_{h,k}(y_j)\}_{j=0,1,\cdots}$ 收敛。

由于密度函数 $\{\hat{f}_{h,k}(y_j)\}_{j=0,1,\cdots}$ 单调递增且有界，所以密度函数收敛。

（2）证明均值移动点列 $\{y_t\}_{t=0,1,\cdots,N}$ 收敛。

①如果存在 j 使得 $y_{j+1}=y_j$，可知移动量为0，所以 y_j 收敛；

②在整个过程中 y_{j+1} 和 y_j 始终不相等。

对于第二种情况的证明如下：

设密度函数迭代 j 步得：

$$m_j=\frac{\sum_{i=1}^{n}(x_i-y_j)g(\|\frac{x_i-y_j}{h}\|^2)}{\sum_{i=1}^{n}g(\|\frac{x_i-y_j}{h}\|^2)}$$

式中 n 为样本个数。

因为 n 有界，$\|x_i-y_j\|$ 有界，$g(\|\cdot\|^2)\geqslant0$ 且有界，所以 $\|m_j\|$ 有界。

由于函数密度随均值移动点单调递增得：

$$\begin{aligned}\hat{f}_{h,k}(y_{j+p})-\hat{f}_{h,k}(y_j)&=\hat{f}_{h,k}(y_{j+p})-\hat{f}_{h,k}(y_{j+p-1})+\hat{f}_{h,k}(y_{j+p-1})-\\&\hat{f}_{h,k}(y_{j+p-2})+\cdots+\hat{f}_{h,k}(y_{j+1})-\hat{f}_{h,k}(y_j)\geqslant\\&\sum_{l=0}^{p-1}(\|m_{l+j}\|^2\sum_{i=1}^{N}g(\|\frac{y_{j+l}-x_i}{h}\|^2))\geqslant\\&M\sum_{l=0}^{p-1}(\|y_{l+j+1}-y_{l+j}\|^2)\end{aligned} \tag{7}$$

式中 M 为 $\min\{\sum_{i=1}^{N}g(\|\frac{y_j-x_i}{h}\|^2),\sum_{i=1}^{N}g(\|\frac{y_{j+1}-x_i}{h}\|^2),\cdots,\sum_{i=1}^{N}g(\|\frac{y_{j+l-1}-x_i}{h}\|^2)\}$。

由于 $f_{h,k}(y_j)$ 单调递增且收敛，所以存在 N_1，当 $j>N_1$ 时

$$0<\hat{f}_{h,k}(y_{j+1})-\hat{f}_{h,k}(y_j)\leqslant\varepsilon \tag{8.1}$$

$$0<\hat{f}_{h,k}(y_{j+2})-\hat{f}_{h,k}(y_{j+1})\leqslant\varepsilon \tag{8.2}$$

$$\vdots$$

$$0<\hat{f}_{h,k}(y_{j+p})-\hat{f}_{h,k}(y_{j+p-1})\leqslant\varepsilon \tag{8.p}$$

式中 ε 为任意给定的正数,p 为有界整数(由于密度函数收敛)。

式(8.1)到(8.p)求和得:

$$0<\hat{f}_{h,k}(y_{j+p})-\hat{f}_{h,k}(y_j)\leqslant p\varepsilon \tag{9}$$

由于 ε 为任意的正数,取 $\varepsilon=\frac{1}{p\cdot N}(N\to\infty)$

式(9)得:

$$0<\hat{f}_{h,k}(y_{j+p})-\hat{f}_{h,k}(y_j)\leqslant\frac{1}{N}$$

设 $\eta=1/N$,上式变为:

$$0<\hat{f}_{h,k}(y_{j+p})-\hat{f}_{h,k}(y_j)\leqslant\eta \tag{10}$$

结合式(7)得:

$$M\sum_{l=0}^{p-1}(\|y_{l+j+1}-y_{l+j}\|^2)\leqslant\eta$$

因为 M 为正值且有界,所以

$$\sum_{l=0}^{p-1}(\|y_{l+j+1}-y_{l+j}\|^2)\leqslant\frac{\eta}{M}$$

因为 η 为任意的小,当 $\eta\to0$ 时,$\sum_{l=0}^{p-1}(\|y_{l+j+1}-y_{l+j}\|^2)\to0$,所以得:

$$\|y_{j+1}-y_j\|^2+\|y_{j+2}-y_{j+1}\|^2+\cdots+\|y_{j+p}-y_{j+p-1}\|^2\to0$$

上式中每一项都大于等于0,所以式中每一项都趋近于0。则存在一个任意小的正数 δ,得:

$$0\leqslant\|y_{j+i}-y_{j+i-1}\|^2_{i=0,1,\cdots,p}\leqslant\delta$$

所以 $0\leqslant\|y_{j+i}-y_{j+i-1}\|_{i=0,1,\cdots,p}\leqslant\sqrt{\delta}$。

$$0\leqslant\sum_{l=0}^{p-1}(\|y_{j+l+1}-y_{l+j}\|)\leqslant p\sqrt{\delta}$$

取 $\sqrt{\delta}=\frac{1}{pN}(N\to\infty)$ 为正的任意小,得:

$$\sum_{l=0}^{p-1}(\|y_{l+j+1}-y_{l+j}\|)\to0 \tag{11}$$

由

$$\sum_{l=0}^{p-1}\|y_{l+j+1}-y_{l+j}\|=\|y_{j+1}-y_j\|+\|y_{j+2}-y_{j+1}\|+\cdots+\|y_{j+p}-y_{j+p-1}\| \tag{12}$$

根据距离不等式性质得:

$$\|y_{j+1}-y_j\|+\|y_{j+2}-y_{j+1}\|\geqslant\|y_{j+2}-y_j\|$$

$$\|y_{j+2}-y_j\|+\|y_{j+3}-y_{j+2}\|\geqslant\|y_{j+3}-y_j\|$$

$$\vdots$$

$$\|y_{j+p-1}-y_j\|+\|y_{j+p}-y_{j+p-1}\|\geqslant\|y_{j+p}-y_j\|$$

对上式两侧求和得:

$$\sum_{l=0}^{p-1}\|y_{l+j+1}-y_{l+j}\|\geqslant\|y_{j+p}-y_j\|$$

结合式(7)得:

$$\|y_{j+p}-y_j\|\to0$$

由 $y_{j+p}=y_{j+p-1}+m_{j+p-1}$ 递推得到:

$$y_{j+p}=y_j+m_{j+p-1}+m_{j+p-2}+\cdots+m_j\Rightarrow y_{j+p}-y_j=m_{j+p-1}+m_{j+p-2}+\cdots+m_j$$

$$\|y_{j+p}-y_j\|=\|m_{j+p-1}+m_{j+p-2}+\cdots+m_j\| \tag{13}$$

所以存在任意的 λ,使得

$$\|m_{j+p-1}+m_{j+p-2}+\cdots+m_j\|\leqslant\lambda$$

存在自然数 N_1,当 $j>N_1$ 时,对于任意的自然数 P 都有 $\|m_{j+p-1}+m_{j+p-2}+\cdots+m_j\|\leqslant\lambda$。所以根据柯西审敛定理和 y_1 为已知值,知 $y_j=y_1+\sum_{i=1}^{j-1}m_i$ 收敛。

证毕。

定理 2 当轮廓函数 $k(x)$ 为单调递减的凸函数时,同一窗内连续均值移动矢量之间夹角不大于90°。

证明

设 y_j,y_{j+1} 和 y_{j+2} 分别为均值移动点列中连续三个点。不失一般性设 $y_j=0$。

$$y_{j+1}=\alpha\frac{\sum_{i=1}^{n}x_i g\left(\left\|\frac{x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n}g\left(\left\|\frac{x_i}{h}\right\|^2\right)} \tag{14}$$

其中 $\alpha\in(0,1]$ 为步长变化系数。

x 取值空间 R^d 由以下三个空间组成:

$$D_1=\{x\in R^d|\alpha y_{j+1}^{\mathrm{T}}x\leqslant\frac{1}{2}\|y_{j+1}\|^2\}$$

$$D_2=\{x\in R^d|\frac{1}{2}\|y_{j+1}\|^2\leqslant\alpha y_{j+1}^{\mathrm{T}}x\leqslant\|y_{j+1}\|^2\}$$

$$D_3=\{x\in R^d|\|y_{j+1}\|^2\leqslant\alpha y_{j+1}^{\mathrm{T}}x\}$$

根据公式(14)得:

$$\sum_{i=1}^{n}y_{j+1}g\left(\left\|\frac{x_i}{h}\right\|^2\right)-\alpha\sum_{i=1}^{n}x_i g\left(\left\|\frac{x_i}{h}\right\|^2\right)=0$$

$$\sum_{i=1}^{n}(y_{j+1}-\alpha x_i)g\left(\left\|\frac{x_i}{h}\right\|^2\right)=0$$

上式两侧同时乘以 y_{j+1}^{T} 得:

$$\sum_{i=1}^{n}(\|y_{j+1}\|^2-\alpha y_{j+1}^{\mathrm{T}}x_i)g\left(\left\|\frac{x_i}{h}\right\|^2\right)=0$$

$$\Rightarrow\sum_{i\in(D_1\cup D_2\cup D_3)}(\|y_{j+1}\|^2-\alpha y_{j+1}^{\mathrm{T}}x_i)g\left(\left\|\frac{x_i}{h}\right\|^2\right)=0$$

所以上式可变为:

$$\sum_{x_i\in D_2}(\|y_{j+1}\|^2-\alpha y_{j+1}^{\mathrm{T}}x_i)g\left(\left\|\frac{x_i}{h}\right\|^2\right)=$$

$$\sum_{x_i\in(D_1+D_3)}(\alpha y_{j+1}^{\mathrm{T}}x_i-\|y_{j+1}\|^2)g\left(\left\|\frac{x_i}{h}\right\|^2\right) \tag{15}$$

在 D_2 中 $\|y_{j+1}\|^2-\alpha y_{j+1}^{\mathrm{T}}x_i\geqslant0$ 且 $\frac{1}{2}\|y_{j+1}\|^2-\alpha y_{j+1}^{\mathrm{T}}x_i\leqslant0$ 得:

$$\|y_{j+1}-x_i\|^2\geqslant\|x_i\|^2+\left(1-\frac{2}{\alpha}\right)\|y_{j+1}\|^2$$

由于 $\|x_i\|^2+\left(1-\frac{2}{\alpha}\right)\|y_{j+1}\|^2$ 不一定大于零。

$$设\ p=\begin{cases}\dfrac{\|x_i\|^2+\left(1-\frac{2}{\alpha}\right)\|y_{j+1}\|^2}{h^2} & \text{if }\left(\|x_i\|^2+\left(1-\frac{2}{\alpha}\right)\|y_{j+1}\|^2\right)\geqslant0\\ 0 & \text{else}\end{cases}$$

由于 $k(x)$ 为单调递减的凸函数。

$$g(\|x\|^2)=-k'(\|x\|^2)>0\ 且\ g'(\|x\|^2)=-k''(\|x\|^2)\leqslant0$$

所以 $g(\|x\|^2)$ 大于0且单调递减。

所以 $\sum_{x_i\in D_2}(\|y_{j+1}\|^2-\alpha y_{j+1}^{\mathrm{T}}x_i)g\left(\left\|\frac{y_{j+1}-x_i}{h}\right\|^2\right)\leqslant\sum_{x_i\in D_2}(\|y_{j+1}\|^2-$

$\alpha y_{j+1}^{\mathrm{T}}x_i)g(p)$

根据公式(15)知：

$$\sum_{x_i\in D_2}(\|y_{j+1}\|^2-\alpha y_{j+1}^{\mathrm{T}}x_i)g(\|\frac{y_{j+1}-x_i}{h}\|^2)\leqslant\sum_{x_i\in D_1\cup D_3}(\alpha y_{j+1}^{\mathrm{T}}x_i-\|y_{j+1}\|^2)g(p)$$

在上式两侧加式：

$$\sum_{x_i\in D_1\cup D_3}(\|y_{j+1}\|^2-\alpha y_{j+1}^{\mathrm{T}}x_i)g(\|\frac{y_{j+1}-x_i}{h}\|^2)$$

得：$\sum_{x_i\in D_1\cup D_2\cup D_3}(\|y_{j+1}\|^2-\alpha y_{j+1}^{\mathrm{T}}x_i)g(\|\frac{y_{j+1}-x_i}{h}\|^2)\leqslant$

$$\sum_{x_i\in D_1\cup D_3}(\|y_{j+1}\|^2-\alpha y_{j+1}^{\mathrm{T}}x_i)(g(\|\frac{y_{j+1}-x_i}{h}\|^2)-g(p))\quad(16)$$

在 D_1 中 $\alpha y_{j+1}^{\mathrm{T}}x\leqslant\frac{1}{2}\|y_{j+1}\|^2$

$$\|y_{j+1}-x_i\|^2=\|y_{j+1}\|^2+\|x_i\|^2-2y_{j+1}^{\mathrm{T}}x_i\geqslant\|x_i\|^2+(1-\frac{1}{\alpha})\|y_{j+1}\|^2$$

讨论 $\sum_{x_i\in D_1}(\|y_{j+1}\|^2-\alpha y_{j+1}^{\mathrm{T}}x_i)(g(\|\frac{y_{j+1}-x_i}{h}\|^2)-g(p))$ 是否小于 0。

因为 $0\leqslant p<\|x_i\|^2+(1-\frac{1}{\alpha})\|y_{j+1}\|^2$ 且函数 $g(\|x\|^2)$ 单调递减得：

$$g(\|\frac{y_{j+1}-x_i}{h}\|^2)\leqslant g(p)$$

知：

$$\sum_{x_i\in D_1}(\|y_{j+1}\|^2-\alpha y_{j+1}^{\mathrm{T}}x_i)(g(\|\frac{y_{j+1}-x_i}{h}\|^2)-g(p))\leqslant 0\quad(17)$$

在 D_3 中 $\|y_{j+1}\|^2-\alpha y_{j+1}^{\mathrm{T}}x\leqslant 0$。

$$0\leqslant\|y_{j+1}-x_i\|^2=$$

$$\|y_{j+1}\|^2+\|x_i\|^2-2y_{j+1}^{\mathrm{T}}x_i\leqslant\|x_i\|^2+(1-\frac{2}{\alpha})\|y_{j+1}\|^2$$

讨论 $\sum_{x_i\in D_3}(\|y_{j+1}\|^2-\alpha y_{j+1}^{\mathrm{T}}x_i)(g(\|\frac{y_{j+1}-x_i}{h}\|^2)-g(p))$ 是否小于零。

当 $p=0$ 时，即 $\|y_{j+1}-x_i\|^2=0$，此时 $g(\|\frac{y_{j+1}-x_i}{h}\|^2)=g(p)$。

当 $p=\frac{\|x_i\|^2+(1-\frac{2}{\alpha})\|y_{j+1}\|^2}{h^2}$ 时，$g(\|\frac{y_{j+1}-x_i}{h}\|^2)\geqslant g(p)$。

于是

$$\sum_{x_i\in D_3}(\|y_{j+1}\|^2-\alpha y_{j+1}^{\mathrm{T}}x_i)(g(\|\frac{y_{j+1}-x_i}{h}\|^2)-g(p))\leqslant 0\quad(18)$$

根据式(17)和(18)得：

$$\sum_{x_i\in D_1\cup D_3}(\|y_{j+1}\|^2-\alpha y_{j+1}^{\mathrm{T}}x_i)(g(\|\frac{y_{j+1}-x_i}{h}\|^2)-g(p))\leqslant 0$$

结合式(16)得：

$$\sum_{x_i\in D_1\cup D_2\cup D_3}(\|y_{j+1}\|^2-\alpha y_{j+1}^{\mathrm{T}}x_i)g(\|\frac{y_{j+1}-x_i}{h}\|^2)\leqslant 0$$

因为 $g(\|x\|^2)=-k'(\|x\|^2)>0$，所以

$$\frac{y_{j+1}^{\mathrm{T}}\sum_{x_i\in D_1\cup D_2\cup D_3}y_{j+1}g(\|\frac{y_{j+1}-x_i}{h}\|^2)}{\sum_{x_i\in D_1\cup D_2\cup D_3}g(\|\frac{y_{j+1}-x_i}{h}\|^2)}\leqslant$$

$$\frac{y_{j+1}^{\mathrm{T}}\cdot\alpha\sum_{x_i\in D_1\cup D_2\cup D_3}x_i g(\|\frac{y_{j+1}-x_i}{h}\|^2)}{\sum_{x_i\in D_1\cup D_2\cup D_3}g(\|\frac{y_{j+1}-x_i}{h}\|^2)}$$

$$\Rightarrow\|y_{j+1}\|^2\leqslant y_{j+1}^{\mathrm{T}}y_{j+2}\Rightarrow\frac{y_{j+1}^{\mathrm{T}}(y_{j+2}-y_{j+1})}{\|y_{j+1}\|\,\|y_{j+2}-y_{j+1}\|}\geqslant 0$$

$$\overset{y_j=0}{\Rightarrow}\frac{m^{\mathrm{T}}(j)m(j+1)}{\|m(j)\|\,\|m(j+1)\|}\geqslant 0$$

由于 $\frac{m^{\mathrm{T}}(j)m(j+1)}{\|m(j)\|\,\|m(j+1)\|}$ 表示均值移动矢量 $m(j)$ 和 $m(j+1)$ 的夹角余弦，所以两连续均值移动矢量的夹角不大于 90°。

3　结论

根据柯西收敛定理很好地证明了均值移动算法的收敛性，解决了以往文献证明的不足。证明了连续均值移动矢量之间的夹角不大于 90°，为均值移动算法的收敛速度等原理的发展奠定了一定的理论基础。

参考文献：

[1] Fukunaga K, Hostetler L D. The estimation of the gradient of a density function, with applications in pattern recognition[J]. IEEE Trans Information Theory, 1975, 21: 32-40.

[2] Cheng Y. Mean shift, mode seeking, and clustering[J]. IEEE Trans Pattern Analysis and Machine Intelligence, 1995, 17(8): 790-799.

[3] Yang C, Duraiswami R, DeMenthon D, et al. Mean-shift analysis using Quasi-Newton methods[C]//IEEE International Conference on Image Processing, 2003: 447-450.

[4] Comaniciu D, Ramesh V, Meer P. Real-time tracking of non-rigid objects using mean shift[C]//Proc of the IEEE Conf on Computer Vision and Pattern Recognition (CVPR), 2000: 142-149.

[5] Fashing M, Tomasi C. Mean shift is a bound optimization[J]. IEEE Trans on Pattern Anal Machine Intell, 2005, 25(3): 471-474.

[6] 李乡儒，吴福朝，吴战义.均值漂移算法的收敛性[J].软件学报，2005，16(3)：365-374.

[7] Comaniciu D, Ramesh V. Mean shift and optimal prediction for efficient object tracking[C]//Mojsilovic A, Hu J. Proc of the IEEE Int'l Conf on Image Processing (ICIP), 2000: 70-73.

[8] Comaniciu D, Ramesh V, Meer P. Kernel-Based object tracking[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564-575.

[9] Zhou XS, Comaniciu D, Krishnan S. An information fusion framework for robust shape tracking[J]. Workshop on Statistical and Computational Theories of Vision SCTV, 2003: 1-24.

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

[11] Comaniciu D, Meer P. Mean shift analysis and applications[C]//Proc of the IEEE Int'l Conf on Computer Vision (ICCV), 1999: 1197-1203.

[12] Barash D, Comaniciu D. A common framework for nonlinear diffusion, adaptive smoothing, bilateral filtering and mean shift[J]. Image and Video Computing, 2003, 22(1): 73-81.

[13] Comaniciu D,Ramesh V,Meer P.The variable band width mean shift and data-driven scale selection[C]//Proc of the IEEE Int'l Conf on Computer Vision(ICCV),2001:438-445.
[14] Comaniciu D.An algorithm for data-driven bandwidth selection[J].IEEE Trans on Pattern Analysis and Machine Intelligence,2003,25(2):281-288.
[15] Comaniciu D,Ramesh V,Bue A D.Multivariate saddle point detection for statistical clustering[C]//Proc of the European Conf Computer Vision(ECCV),2002:561-576.
[16] Bradski G R.Computer vision face tracking for use in a perceptual user interface[J].IEEE Workshop on Applications of Computer Vision.Stoughton:Printing House,1998:214-219.
[17] Collins R T.Mean-Shift blob tracking through scale space [C]//Proc of the Conf.on Computer Vision and Pattern Recognition(CVPR),2003:18-20.
[18] Yang C,Duraiswami R,Davis L.Similarity measure for nonparametric kernel density based object tracking[C]//8th Annual Conference Neural Information Processing Systems,2005:1561-1568.
[19] Yang C,Duraiswami R,Elgammal A,et al.Real-time kernel-based tracking in joint feature-spatial spaces,Tech Rep CS-TR-4567[R].University of Maryland,2003.
[20] 朱胜利,朱善安,李旭超.快速运动目标的 Mean shift 跟踪算法[J].光电工程,2006,33(5):66-70.
[21] 贾静平,张艳宁,柴艳妹,等.目标多自由度 Mean Shift 序列图像跟踪算法[J].西北工业大学学报,2005,23(5):618-622.
[22] 李培华.一种改进的 mean shift 跟踪算法[J].自动化学报,2007,33(4):347-354.
[23] 程伟,杨杰.一种基于均值移位的红外目标跟踪新方法[J].红外与毫米波学报,2005,24(3):231-235.
[24] Deguchi K,Kawanaka O,Okatani T.Object tracking by the mean-shift of regional colour distribution combined with the particle-filter algorithm[C]//Proc ICPR,2004:506-509.
[25] Bandera A,Perez-Lorenzo J M.Mean shift based clustering of Hough domain for fast line segment detection[J].Pattern Recognition Letters,2006,27:578-586.
[26] Wang Yuzhong.Unsupervised color-texture segmentation based on soft criterionwith adaptive mean-shift clustering[J].PRL,2006,27(5):386-392.
[27] Wang J,Thiesson B,Xu Y Q,et al.Image and video segmentation by anisotropic kernel mean shift[C]//Proc of the European Conf on Computer Vision(ECCV),2004.
[28] Carreira-Perpiñán M Á.Fast nonparametric clustering with Gaussian blurring mean-shift[C]//Proc 23rd Int'l Conf Machine Learning,2006.
[29] Georgescu,B.Shimshoni,Meer P.Mean shift based clustering in high dimensions:A texture classification example[C]//IEEE Int Conf on Computer Vision,2003(2):456-463.
[30] Tek H,Comaniciu D,Williams J.Vessel detection by mean-shift based ray propagation[C]//IEEE Workshop on Mathematical Methods in Biomedical Image Analysis,Hawaii,2001.
[31] 刘蓉,段福庆.基于均值漂移的自适应滤波及其在光谱信号处理中的应用[J].电子与信息学报,2006,28(2):312-316.
[32] Comaniciu D,Ramesh V,Meer P.Real-Time tracking of non-rigid objects using mean shift[C]//Proc of the IEEE Conf on Computer Vision and Pattern Recognition(CVPR),2000:142-149.
[33] Comaniciu D,Meer P.Mean shift:A robust approach toward feature space analysis[J].IEEE Trans on Pattern Analysis and Machine Intelligence,2002,24(5):603-619.
[34] 同济大学数学教研室.高等数学[M].北京:高等教育出版社,1996.

(上接33页)

表3 多调控基因的隶属度分布

基因	调控因子	所属组别	在各个组别中的隶属度				
			C1	C2	C3	C4	C5
YBR020W	GAL4 MIG1 REB1	1,5	0.422	0.004	0.005	0.002	0.567
YCL040W	GCR1 MSN2 MSN4	4,5	0	0.012	0.043	0.127	0.818
YHR030C	RLM1 SWI4 SWI6	3,4	0.01	0.073	0.293	0.619	0.005
YJR094C	IME4 RME1	2,3	0	0.448	0.450	0.102	0

YHR030C,YJR094C 也分别在其隶属度上体现了其多个调控因子共同作用的影响。满足上述条件(2),说明该文方法得到的模糊聚类隶属度可观察到实际存在的多调控关系。

5 结论

通过对实验结果的分析,TransChisq 相关系数能够准确发现基因表达数据中的共调控基因,改进的模糊 C 均值聚类算法能够将基因按照其生物实际基因功能准确分类,并体现多调控基因的从属关系,为从基因表达数据中挖掘更符合实际生物意义的信息提供了更有力的方法。

参考文献:

[1] Agrawal R,Gehrke J,Gunopulos D,et al.Automatic subspace clustering of high dimensional data for data mining applications[C]//Proceedings of ACM SIGMODs 99 International Conference.Management of Data(Philadelphia,PA),1998:94-105.
[2] Alizadeh A,Eisen M B,Davis R E,et al.Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling[J].Nature,2000,403:503-511.
[3] Hsu A L,Tang S,Halgamuge S K.An unsupervised hierarchical dynamic self-organizing approach to cancer class discovery and marker gene identification in microarray data[J].Bioinformatics,2003,19:2131-2140.
[4] Krishna K,Narasimha M M.Genetic K-means algorithm[J].IEEE Trans Syst Man Cybern:Part B,1999,29:433-439.
[5] Kim K,Zhang Shibo,Jiang Keni,et al.Measuring similarities between gene expression profiles through new data transformations[J].BMC Bioinformatics,2007,8.
[6] Spellman P T,Sherlock G,Zhang M Q,et al.Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization[J].Mol Biol Cell,1998,9:3273-3297.
[7] Lee T I,Rinaldi N J,Robert F,et al.Transcriptional regulatory networks in Saccharomyces cerevisiae[J].Science,2002,298:799-804.