

12

67-72

Fuzzy C-Means 算法中隶属度信息 在特征空间的分布特性分析及改进方法

胡世英 周源华

(上海交通大学图像通信与信息处理研究所, 上海, 200030)

C235
TP391.41

摘要 首先推导了 Fuzzy C-Means 算法在特征空间的迭代公式, 然后就其隶属度信息在特征空间的分布缺陷提出两种改进方法: 一是通过引入选择注意性参数控制隶属度信息的分布; 二是从条件概率出发构造类置信度取代原隶属度. 实验表明这两种方法均起到了较好的效果.

关键词 Fuzzy C-Means 算法, 隶属度, 选择注意性参数, 置信度.

FCM算法 模糊聚类 图像

引言

Fuzzy C-Means (FCM)算法是以类内均方误差和为准则函数的模糊聚类算法, 该方法由 Dunn^[1]提出, 经 Bezdek^[2]推广后获得了广泛的应用^[3~6]. 与其它分割算法不同, FCM 算法不仅为后续处理提供分割结果, 而且还提供相对于各类的隶属程度信息. 但原有算法提供的隶属度信息在特征空间的分布通常会与图像本身的特征分布不符, 为此本文提出两种改进方法: (1) 将选择注意性参数引入准则函数, 根据先验知识控制隶属度信息的分布; (2) 从条件概率的角度出发构造类置信度取代原有的隶属度信息.

1 特征空间中的 FCM 算法

FCM 算法通过迭代优化基于类内均方误差和这一准则函数实现对输入数据的优化分类, 文献[2]中对其迭代过程及算法特性进行了详细的论述. 传统的 FCM 方法是在样本空间中进行的, 下面将针对特征空间进行讨论.

令 Z 为样本集 X 的特征集, z 为 Z 中任一特征. 设 $P(z)$ 为特征 z 出现的概率, $P(z) \cong k/N$, 其中 N 为样本总数, k 为特征 z 出现的频率, 则 FCM 算法在特征空间中可表述为

$$\text{minimize } J_m(U, V) = \sum_{z \in Z} \sum_{i=1}^C P(z) u_i^m(z) (d_i(z))^2, \quad (1)$$

$$\sum_{i=1}^C u_i(z) = 1, \quad z \in Z \quad (2)$$

$$u_i(z) \geq 0, \quad z \in Z, 1 \leq i \leq C \quad (3)$$

其中 C 为预定的类数, $m > 1$ 为模糊指数, $u_i(z)$ 表示特征 z 相对于第 i 类的隶属度, $d_i(z) = \|z - v_i\|$, $z, v_i \in R^p$, p 为特征空间维数, $\|\cdot\|$ 是 R^p 中任何一种由内积诱导的范数, $U =$

$\{u_i(z)\}$ 表示隶属度矩阵, $V = \{v_1, \dots, v_c\}$ 表示类均值矩阵. 利用与文献[2]相似的方法可以证明使准则函数最小化的必要条件为

$$u_i(z) = \left[\sum_{j=1}^c (d_i(z)/d_j(z))^{2/(m-1)} \right]^{-1}, \quad (4)$$

若 $d_i(z) = 0$, 则 $u_i(z) = 1$ 且 $u_i(z)$ 对 $j \neq i$.

$$v_i = \sum_{z \in Z} u_i^m(z) z P(z) / \sum_{z \in Z} u_i^m(z) P(z), \quad (5)$$

由于样本数 $|X| \geq$ 样本特征数 $|Z|$, 因而在特征空间计算可以在一定程度上减少运算量. 尤其将 FCM 用于图像处理时, 由于图像像素间的相关性, 在特征空间计算的优势更加明显. 下面的讨论均在特征空间中进行.

2 隶属度信息分布特性的改进方法

2.1 隶属度信息分布特性分析

下面以标准图像 Lenna 脸部的灰度直方图为例进行讨论. 参数设置如下: $C=3, m=2, E=1e-4$, 随机初始化 $\{v_1, v_2, v_3\}$. 采用欧氏距离, 迭代得到最终的 3 个聚类中心为: $v_1=31, v_2=115, v_3=181$. 按式(4)可以画出各类的隶属度分布曲线示意图, 如图 1(a). 由图可见, FCM 的隶属度曲线存在两点问题. (1) 隶属度曲线的形状与图像本身的特征分布不符, 图 1(a)中 A、B 两类的特征分布存在明显差异, 但其隶属度曲线形状却很相近, 将这种隶属度信息用作后续处理是有偏差的; (2) 隶属度曲线存在旁瓣, 模糊集 $\{u_i(z)\} \{z \in Z\}$ 为非凸模糊集^[5], 这与隶属度所要求表示的以离中心为衡量标准的可能性度量有差距, 故应采用隶属度函数的截函数以消除旁瓣影响. 下面的讨论均采用消除旁瓣后的隶属度值.

2.2 选择注意性参数的运用

FCM 方法中的隶属度曲线与特征分布不符主要是由于该算法是一种无先验知识参与的盲目的聚类方法. 这里将根据特征分布的先验知识构造选择注意性参数项 $a(i, z)$, 并将其加入到准则函数中, 以控制隶属度曲线的形状. 通常可令 $\partial J_m(i, z) / \partial z$, 则 $a(i, z) \equiv a(i)$, 于是新的准则函数定义为

$$J_m(U, V) = \sum_{z \in Z} \sum_{i=1}^c a(i) P(z) u_i^m(z) (d_i(z))^2. \quad (6)$$

可以证明使上式最小化的必要条件为

$$u_i(z) = \left[\sum_{j=1}^c \left(\frac{a(j)d_j(z)}{a(i)d_i(z)} \right)^{2/(m-1)} \right]^{-1}. \quad (7)$$

若 $d_i(z) = 0$, 则 $u_i(z) = 1$ 且 $u_i(z) = 0$ 对 $j \neq i$.

$$v_i = \sum_{z \in Z} u_i^m(z) z P(z) / \sum_{z \in Z} u_i^m(z) P(z). \quad (8)$$

假定对 Lenna 图有下述先验知识: $C=3, v_1 < v_2 < v_3, C_1$ 比 C_2, C_3 紧密. 则可设置 $a(2) = a(3) = 1, a(1) = 2$. 此时按式(7)得到的隶属度曲线如图 1(b). 比较图 1(a)和(b)可看出, 后者的隶属度曲线更符合相应的特征分布状况. 引入选择注意性参数 $a(i, z)$ 的缺陷在于 $a(i, z)$ 的确定有赖于先验知识, 并且 $a(i, z)$ 的引入使迭代过程对类中心的初始化值更敏感.

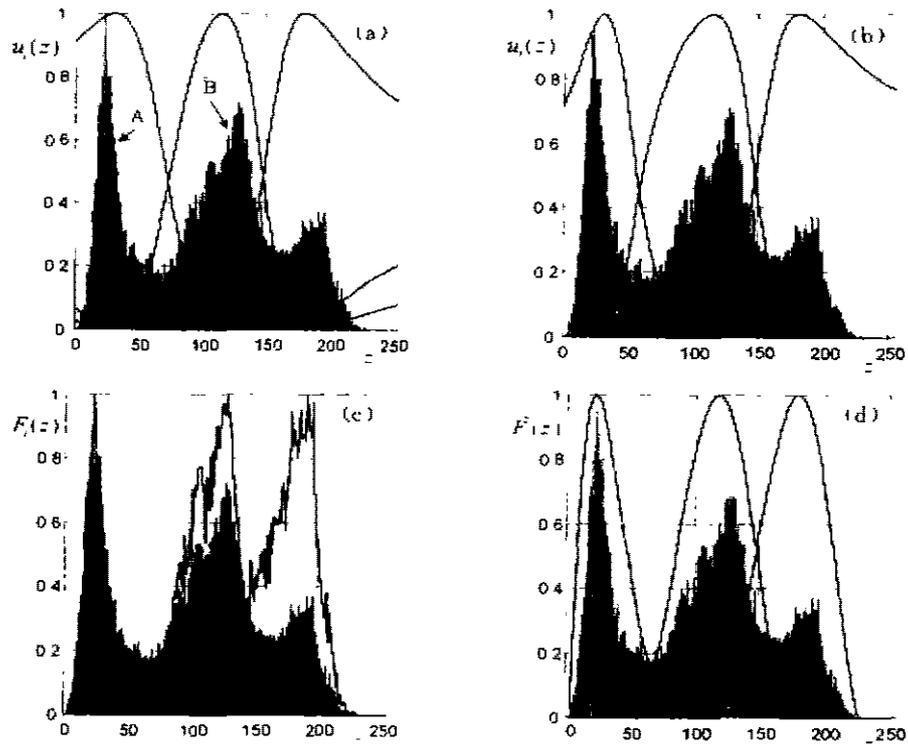


图1 (a)原始FCM算法隶属度曲线图, (b)引入选择性注意性参数后的隶属度曲线图
(c) FCM 原始类置信曲线图, (d) FCM 规范化拟合类置信曲线图

Fig. 1 (a) membership curve diagram of original FCM algorithm, (b) membership curve diagram with selective attention parameters, (c) original fiducial curve diagram of FCM algorithm, (d) normalized fitting fiducial curve diagram of FCM algorithm

2.3 类置信函数定义

一般希望各类的隶属度曲线与相应的特征分布有关,上面选择性注意性参数的引入就是针对这一问题提出的.下面将从另一角度考察这一问题.

定义: FCM 原始类置信函数(Original Fiducial Function, 简称 FCM-OFF) $F_i(z)$ 为

$$F_i(z) = f_i(z)/f_{i,max}, \quad (9)$$

$$f_i(z) = u_i^n(z)p(z), \quad (10)$$

其中 $f_{i,max} = \max_{z \in Z} f_i(z)$, $u_i(z)$ 为 FCM 算法的迭代结果; n 为置信指数, 满足 $n \geq 1$; 要求 $f_i(z)$ 不恒为零. 上述类置信函数将原来的隶属度信息与特征空间分布信息相结合, 从而获得比 $u_i(z)$ 更恰当的对特征空间的描述.

下面从概率的角度分析上述定义的含义. 在数据处理过程中往往需要在限定某一类的条件下得到某个特征值相对于该类的置信度, 这个值可以用条件概率 $p(z|C_i)$ 来表示. 对于

FCM 中定义的隶属度 $u_i(z)$, 其约束条件与条件概率的定义 $\sum_{i=1}^C p(C_i|z) = 1$ 一致, 因而可以把它看成后验概率 $p(C_i|z)$. 在上述类置信函数的定义中, 当 $n=1$ 时, $f_i(z)$ 与联合概率密度 $p(C_i|z)$ 相当, 因而类置信函数间接提供了条件概率 $p(z|C_i)$ 的信息.

对 Lenna 图, 设定 $C=3, m=2$, 随机初始化类中心, 得到的各类置信函数曲线如图 1(c) 所示. 从图中可以看出, 置信曲线虽能符合特征分布, 但存在较大起伏, 这主要是由于仅从一幅图得到的 $P(z)$ 只是特征分布的一种近似表示, 带有较大随机噪声, 这对后续处理是不利的, 故采用拟合的方法平滑曲线.

定义: FCM 拟合类置信函数 (Fitting Fiducial Function, 简称 FCM-FFF) $F_i(z)$ 为

$$\|\hat{F}_i - F_i\| = \min_{G \in L} \|F_i - G\|. \quad (11)$$

式中 $\hat{F}_i(z)$ 为某范数意义下的对 $F_i(z)$ 的最佳拟合, 其中, L 表示满足一定可微条件的拟合函数族的子集.

FCM-FFF 的引入在一定程度上消除了输入数据的随机噪声. 从这个意义上看, 在实现拟合过程时应注意以下准则: (1) 拟合函数族应满足高阶可微, 以保证曲线的平滑性要求; (2) 函数子集的选择应防止出现过多的拐点. 以图 1(c) 为例, 采用多项式拟合时, 其指数不应过大, 这里取为 8. 为了消除拟合对 FCM-FFF 幅值的影响, 对其进行归一化处理.

定义: FCM 规范化拟合类置信函数 (Normalized Fitting Fiducial Function, 简称 FCM-NFFF) $\bar{F}_i(z)$ 为

$$\bar{F}_i(z) = |\hat{F}_i(z) / \min_{z \in Z} F_i(z)|. \quad (12)$$

考虑到拟合过程对类中心的影响, 重新计算类中心为

$$\tilde{v}_i = \sum_{z \in Z} \bar{F}_i(z) z / \sum_{z \in Z} \bar{F}_i(z). \quad (13)$$

根据上述讨论得到 FCM 规范化拟合类置信函数 $\bar{F}_i(z)$ 如图 1(d) 所示, 显然类置信函数比隶属度对特征空间的描述更恰当.

3 实验结果

图 1 和 2 给出了对标准图像 Lenna 脸部的灰度直方图及一车牌图片的直方图: (该直方图在 255 级灰度处有 1 类) 采用原始 FCM 方法和两种改进方法得到的曲线图, z 表示灰度级, $u_i(z)$, $F_i(z)$, $\bar{F}_i(z)$ 分别表示灰度级 z 相对与第 i ($1 \leq i \leq C$) 类的隶属度值、原始类置信度值和规范化拟合类置信度值 (为便于比较, 每幅图中都同时显示规范化后的直方图). 从图中可以看出引入选择注意性参数及定义类置信度两种方法从不同程度上改善了隶属度信息在特征空间中的分布特性, 从而为后续处理提供了更为准确的信息. 表 1 列出了有关参数和实验数据 (迭代参数 $C=3, m=2$, 随机初始化类中心), 从中可以看出由于选择注意性参数体现了类分布的先验知识, 因而对类中心的定位更加准确, 而 FCM-NFFF 方法则更侧重于体现 $P(z)$ 的信息. 另外, 由于消除了旁瓣的影响, 也使后两种方法与原始 FCM 方法对中心的定位有差别.

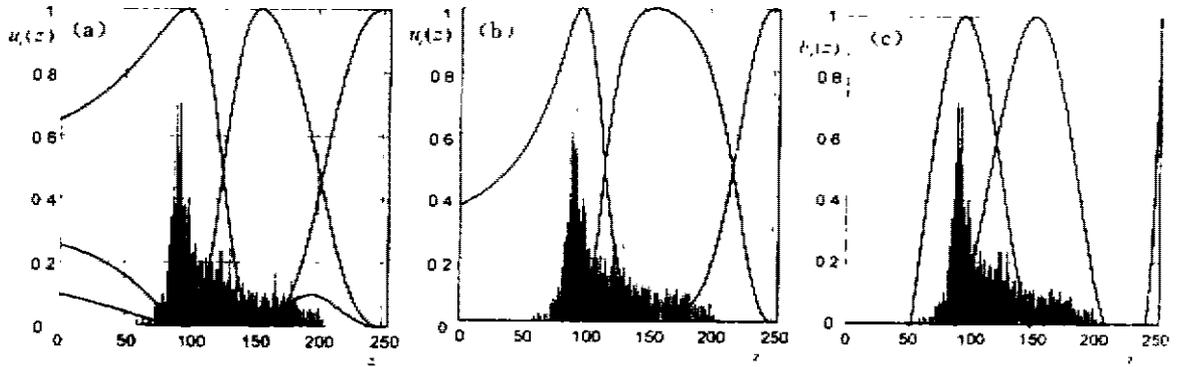


图2 (a) 原始FCM隶属度曲线图, (b) 引入选择注意项后得到的隶属度曲线图, (c) FCM规范化拟合类置信曲线图

Fig. 2 (a) membership curve diagram of original FCM algorithm, (b) membership curve diagram with selective attention parameters, (c) normalized fitting fiducial curve diagram of FCM algorithm

表1 实验参数及结果

Table 1 Experimental parameters and results

		Lenna	车牌
原始FCM方法	分类结果	$v_1=31, v_2=115, v_3=181$	$v_1=97, v_2=155, v_3=250$
引入选择注意项 $a(z)$ 后的FCM方法	$a(z)$ 设置	$a(1)=2, a(2)=a(3)=1$	$a(1)=2, a(2)=1, a(3)=2$
	分类结果	$v_1=31, v_2=115, v_3=181$	$v_1=98, v_2=161, v_3=255$
FCM-NFFF方法	参数设置	置信指数 $n=1$, 多项式拟合指数 $p=8$	
	分类结果	$v_1=35, v_2=116, v_3=177$	$v_1=98, v_2=152, v_3=252$

4 结论

本文在特征空间中考察 FCM 算法, 推导其在特征空间的迭代公式, 通过分析隶属度信息在特征空间中的分布特性, 就其存在的缺陷, 即与特征分布不符, 提出两种改进方法: (1) 是将选择注意性参数引入准则函数, 根据特征分布的先验知识控制隶属度信息的分布状况; (2) 是从概率的角度出发, 利用隶属度及特征分布的双重信息, 定义类置信函数取代隶属度. 实验结果表明这两种方法均起到了较好的效果. 本文的讨论主要以一维特征空间为例, 但以上的结论完全可以推广到高维情况.

REFERENCES

- 1 Dunn J C. *Cybern.*, 1974, 3:32
- 2 Bezdek J C. *Pattern Recognition with Fuzzy Objective Function Algorithm*, New York; Plenum Press, 1981
- 3 Huntsberger T L, Jacobs C L, Cannon R L. *Pattern Recognition*, 1985, 18(2):131
- 4 Mohan M Trivedi, Bezdek J C. *IEEE Tran. on Systems, Man, and Cybernetics*, 1986, SMC-16(4):589
- 5 Guo Gui-Rong. *Fuzzy Pattern Recognition*, Hunan; National Defense Science And Technology University Publishing House(郭桂荣. 模糊模式识别, 湖南:国防科技大学出版社), 1993:
- 6 Helene Caillol, Wojciech Pieczynski, Alain Hillion. *IEEE Tran. on Image Processing*, 1997, 6(3):425

**ANALYSIS AND IMPROVEMENT OF DISTRIBUTION PROPERTY OF
MEMBERSHIP INFORMATION IN FEATURE SPACE
BASED ON FUZZY C-MEANS ALGORITHM**

HU Shi-Ying ZHOU Yuan-Hua

(Institute of Image Communication and Information Processing,
Shanghai Jiaotong University, Shanghai 200030, China)

Abstract The iterative formulas of Fuzzy C-Means algorithm in feature space were deduced before two methods were presented to improve the distribution property of membership information. The first was to control the distribution of membership information by introducing selective attention parameters. The other was to construct cluster fiducial values from the point of probability to replace original membership information. Results of the experiments confirm the effectiveness of the two methods.

Key words fuzzy C-means algorithm, membership values, selective attention parameters, fiducial values.