

改进的模糊 C 均值聚类算法

林亚忠¹, 张会奇², 李新², 栾钦波²

(¹解放军第 175 医院/厦门大学附属东南医院, 福建 漳州 363000;

²厦门大学 计算机科学系, 福建 厦门 361005)

【摘要】 由于传统的模糊 C 均值 (fuzzy C-means, FCM) 算法没有考虑像素点的空间邻域信息, 仅涉及像素的单个灰度, 在处理含有噪声的图像时有很大的局限性, 因此分割效果较差。针对 FCM 的缺陷, 提出一种新的改进算法, 该算法引入 Gibbs 随机场, 将 Gibbs 随机场先验概率与像素点隶属度的乘积作为新的像素隶属度。实验表明, 改进后的算法有良好的分割效果, 既可以较为完整地保留图像边界细节, 又能较好地去除图像的噪声。

【关键词】 模糊 C 均值; 聚类; 噪声; Gibbs 随机场

中图分类号: R318.6; TP391

文献标识码: A

doi:10.3969/j.issn.1674-4659.2013.04.0385

Improved Method for Fuzzy C-Means Clustering Algorithm

LIN Yazhong¹, ZHANG Huiqi², LI Xin², LUAN Qinbo²

(¹PLA 175th Hospital / Southeast Hospital Affiliated to Xiamen University, Zhangzhou 363000, China;

²Department of Computer Science, Xiamen University, Xiamen 361005, China)

[Abstract] The traditional fuzzy C-means (FCM) algorithm has great limitations in dealing with the noisy images owing to not considering the spatial information of the pixels and only involving the pixel gray of a single point, so it's poor in segmenting an image. For the defects of the FCM algorithm, a new improved algorithm is proposed in this article, in which a product of Gibbs priori probability and the membership is regarded as the new pixel membership. Experimental results show that the improved algorithm has a good segmentation result; it can retain more complete edge details of image and can remove the image noise more effectively.

[Key words] Fuzzy C-means (FCM); Clustering; Noise; Gibbs random field

0 引言

图像分割是图像处理和计算机视觉研究领域中的重要技术之一, 是指将图像中具有不同涵义、互不交叉的区域分割开来。常用的分割方法有阈值分割、边缘分割、区域分割等。模糊 C 均值 (fuzzy C-means, FCM) 算法^[1-2]就是一种典型的区域分割算法, 它可以很好地保留图像的细节, 但是由于 FCM 算法没有考虑像素点的空间邻域信息^[3], 在处理噪声图像时效果不佳, 文献 [6-10] 中提到的 Gibbs 随机场可以很好地刻画邻域中像素间的相互关系, 本文正是利用 Gibbs 随机场的这一特性, 改进了 FCM 算法中的隶属度表达, 较好地解决了噪声图像的分割精度问题。

收稿日期: 2013-02-19

基金项目: 南京军区重点项目 (项目编号: 11Z023); 福建省自然科学基金项目 (项目编号: 2008J0312)

作者简介: 林亚忠 (1973-), 男, 福建漳州人, 高级工程师, 博士, 硕士研究生导师, 主要研究方向: 医学图像处理、模式识别; 张会奇 (1987-), 女, 河南新乡人, 硕士研究生, 主要研究方向: 计算机图像处理; 李新 (1987-), 女, 山东泰安人, 硕士研究生, 主要研究方向: 计算机图像处理; 栾钦波 (1987-), 男, 山东临沂人, 硕士研究生, 主要研究方向: 计算机图像处理。

1 传统的 FCM 算法

FCM 算法是 Bezdek 在 1981 年提出的, 是在 K-Means 算法的基础上引入模糊隶属度的概念, 将每个给定的像素点用值在 0 到 1 之间的模糊隶属度来确定其属于各个分类的程度, 通过迭代目标函数得到最优的分类结果。目标函数为:

$$J_{FCM}^{(1)} = \sum_{i=1}^n \sum_{j=1}^c u_{ij}^m \|x_i - v_j\|^2 \quad (1)$$

其中 x_i 表示第 i 个像素点的灰度值; u_{ij} 表示第 i 个像素属于第 j 类的隶属度; v_j 是第 j 个聚类中心; m 是一个模糊加权指数 $1 < m < \infty$ 。

利用拉格朗日乘法得到隶属度和聚类中心的迭代公式②和③。

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|x_i - v_j\|^2}{\|x_i - v_k\|^2} \right)^{\frac{2}{m-1}}} \quad (2)$$

$$v_j = \frac{\sum_{i=1}^n u_{ij}^m \cdot x_i}{\sum_{i=1}^n u_{ij}^m} \quad (3)$$

通过上述两个公式的迭代实现对图像的分割。虽然该算法可以较好地保留图像细节，但由于没有考虑到图像邻域中空间的相互关联信息，使得其在分割噪声图像时，存在分割精度问题。而 Gibbs 随机场充分利用局部的邻域关系很好地刻画了图像像素间的相互关系，能够较好地弥补传统算法对像素邻域信息描述的不足。

2 Gibbs 随机场理论及其相关改进方法

2.1 Markov 与 Gibbs 随机场

在 Markov 随机场理论中，图像指标集 S 中的像素点的空间关系是通过邻域系统 $N = \{N_i | i \in S\}$ [4] 描述的。其中邻域 N_i 表示与点 i 相邻的点的集合， N_i 不包括 i 点本身且满足若 $i \in N_j$ 则 $j \in N_i$ 。当随机场 X 满足非负性和 Markov 性， X 为 Markov 随机场。根据 Hammersley-Clifford 定理 [5]，Markov 随机场与 Gibbs 随机场有一致的对应关系，即 Markov 随机场先验概率可以用 Gibbs 先验概率 [6-9] 表示：

$$P(x) = \frac{\exp(-U(x))}{\sum_{x \in X} \exp(-U(x))} \tag{4}$$

$$U(x) = \sum_{c \in C} V_c(x) \tag{5}$$

其中 $U(x)$ 是能量函数， $V_c(x)$ 是势团 c 上的势能函数。在本文方法中根据多级逻辑模型 (MLL) 计算两点势能函数：

$$V_2(x_i - x_j) = \beta[1 - \delta(x_i - x_j)] \tag{6}$$

其中 β 是对两邻域势团中不相等类别标记的惩罚参数，本文中 β 均取为 1。

2.2 结合 Gibbs 随机场的自适应加权 FCM 改进算法 [10]

文献 [10] 针对传统 FCM 算法的缺点提出一种结合 Gibbs 随机场的自适应加权 FCM 改进方法。该方法结合了 Gibbs 随机场和隶属度场两种场的优点，对两种场进行自适应加权，得到像素的改进隶属度更新公式：

$$U^*(x) = w(x) * P(x) + (1 - w(x)) * U(x) \tag{7}$$

其中， $P(x)$ 为像素 x 的 Gibbs 场先验概率， $U(x)$ 为像素 x 的隶属度值， $w(x)$ 为像素 x 的加权项，取值范围为 $[0, 1]$ ，它的值由像素的邻域标准差信息来决定，计算公式为：

$$w(x) = \exp\left(-\frac{\sigma(x)}{(\max_{y \in N_x} (y - \text{mean}(x)) + \varepsilon) * d}\right) \tag{8}$$

公式⑧中， $\sigma(x)$ 为邻域标准差， y 为像素 x 的邻域像素， $\text{mean}(x)$ 为邻域均值， ε 为一个较小的数， d 为调整系数，具体参数设定及实现见文献 [10]。该算法通过迭代公式⑦和公式③得到较之 FCM 算法更好的分割效果。但此方法在处理噪声比较高或较为复杂的噪声图像时，仍有部分噪声存在，有进一步去除的需求。针对以上不足，本文提出一种新的改进算法，用于解决去除噪声和保留图像细节间相互制约的问题。

3 本文改进的 FCM 算法

FCM 算法由于没有充分考虑图像的空间信息，使得在进行噪声图像的分割时精度不够理想。因此，改进算法通过在 FCM 中引入 Gibbs 随机场，并利用 Gibbs 先验概率来进一步优化隶属度函数表达，可以得到满意的分割结果。优化后隶属度表示为：

$$u_{ij} = u_{ij} * p_{ij} \tag{9}$$

其中 u_{ij} 是第 i 个像素点对第 j 类的隶属度， p_{ij} 表示第 i 个像素点被标记为第 j 类的先验概率。

本文改进算法的实现步骤如下：(1) 初始化分类数 c ，模糊加权系数 m ，迭代终止阈值 ε ，初始聚类中心 $V(0)$ ；(2) 利用传统的 FCM 算法做初始分割得到隶属度矩阵 $U = \{u_{ij}\}$ ；(3) 将传统 FCM 的软分割结果转化为硬分割结果，用公式④、⑤、⑥计算得 Gibbs 随机场先验概率 $P = \{p_{ij}\}$ ；(4) 按照公式⑨计算出新的改进的隶属度 u_{ij}' ，并归一化隶属度矩阵；(5) 按照公式③更新聚类中心 V ；(6) 判断是否满足终止条件 $\|V(t+1) - V(t)\| < \varepsilon$ (t 为迭代次数)，满足则退出，否则转到③步。

4 实验分析

4.1 实验一 人工合成图像的分割

实验一是对 256×256 ，灰度分别为 30、100、170、220 的四种区域构成的合成图像利用 matlab 的 `imnoise` 函数分别加入均值为 0、方差不同的高斯白噪声和椒盐噪声后，利用传统 FCM 算法、自适应加权算法、本文改进算法分割的效果比较。图 1 是各算法对加入不同比例的高斯和椒盐噪声合成图像的分割结果，图 2 是各算法处理不同噪声图像的分割正确率折线图，表 1 和表 2 是各算法处理加入不同高斯噪声比和椒盐噪声比的合成图像的分割正确率统计表。本实验是在 Intel Core2 Duo T5800 2.0G CPU、2G DDR2 内存环境，VC++6.0 软件开发平台下完成，实验中设置每个算法的模糊加权系数 $m = 2$ ，分类数 $c = 4$ ，迭代终止阈值 $\varepsilon = 1.5$ 。

本文定义分割正确率如下：分割正确率 = 图像中被正确分类的像素个数/图像像素的总个数。

图 1 中图 a 是人工合成图像，图 b1、b2、b3 是对图 a 分别加入 15%、8% 的高斯白噪声和 25% 的椒盐噪声后的图像，图 c1、c2、c3 和 d1、d2、d3 及 e1、e2、e3 分别是 FCM 算法、自适应加权算法、本文改进算法对图 b1、b2、b3 处理后的分割结果。

从图 1 的分割结果可以看出：FCM 处理后的图像大部分的噪声都没有被去除，这是因为 FCM 算法没有考虑空间邻域信息，导致处理含有噪声的图像非常困难。而加权自适应算法和本文改进算法都考虑了像素的空间邻域信息，比较图 1 中的图 b3、d3、e3 可知，去噪效果得到明显提高。而与自适应加权算法的图 d3 相比，本文改进算法的图 e3 仅存有极少数的几个分散噪声点，大部分像素点都可以得到正确的分割，去噪效果上更加优异。因此，通过对比三种算法的分割结果可知，本文改进算法的分割效果是最优的。

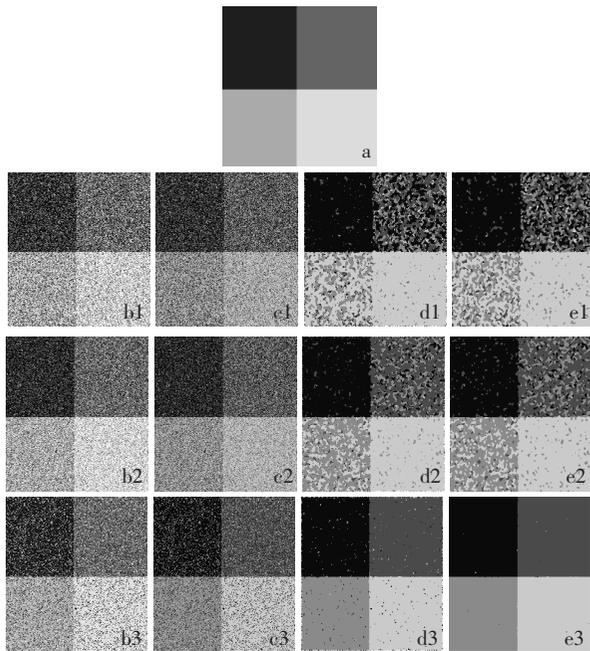


图 1 各算法对加入不同比例的高斯和椒盐噪声合成图像进行分割的结果 (a: 人工合成图像; b1: 加入 15%高斯噪声; c1: FCM 处理结果; d1: 自适应加权处理结果; e1: 本文改进算法结果; b2: 加入 8%高斯噪声; c2: FCM 处理结果; d2: 自适应加权处理结果; e2: 本文改进算法处理结果; b3: 加入 25%椒盐噪声; c3: FCM 处理结果; d3: 自适应加权处理结果; e3: 本文改进算法处理结果)

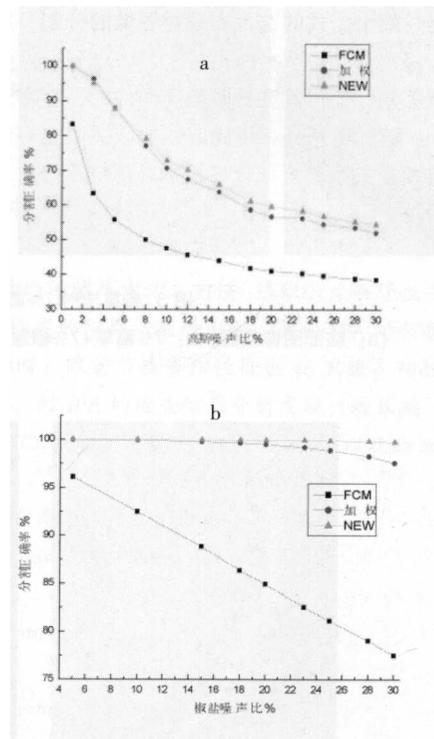


图 2 分割正确率折线图

(a: 各算法处理不同高斯噪声比图像的分割正确率的折线图; b: 各算法处理不同椒盐噪声比图像的分割正确率的折线图)

表 1 各算法处理不同高斯噪声比的合成图像的分割正确率 (%)

项目	高斯 30%	高斯 28%	高斯 25%	高斯 23%	高斯 20%	高斯 18%	高斯 15%	高斯 12%	高斯 10%	高斯 8%	高斯 5%	高斯 3%	高斯 1%
FCM	38.54	38.80	39.60	40.30	40.98	41.82	43.97	45.60	47.54	50.02	55.82	63.33	82.32
自适应加权	52.07	53.52	54.53	56.25	56.73	58.65	63.77	67.47	70.68	77.07	87.97	96.29	99.84
本文改进	54.47	55.31	56.80	58.31	59.56	61.07	65.99	70.16	72.83	79.14	87.69	95.01	99.86

表 2 各算法处理不同椒盐噪声比的合成图像的分割正确率 (%)

项目	椒盐 30%	椒盐 28%	椒盐 25%	椒盐 23%	椒盐 20%	椒盐 18%	椒盐 15%	椒盐 10%	椒盐 5%
FCM	77.53	79.03	81.10	82.53	84.95	86.36	88.85	92.50	96.14
自适应加权	97.57	98.23	98.87	99.19	99.49	99.64	99.78	99.93	99.98
本文改进	99.72	99.81	99.87	99.93	99.93	99.96	99.97	99.99	99.99

从表 1 和表 2 中可以看出：随着噪声比例的不断增多，三种算法对图像分割的正确率都呈现出逐渐递减的趋势。同时，在相同类型、相同比例的噪声下，本文算法的分割正确率是最高的。从图 2a 中看到，虽然当高斯噪声比为 3%时自适应加权算法正确率略高于本文改进算法，但是随着噪声比的不断增加，本文算法的正确率优势逐渐明显，平均正确率比 FCM 高出约 22%，比自适应加权算法高出约 2%。在稳定性上，本文改进算法的效果也相当优异，尤其在图 2b 椒盐噪声折线图中，本文方法的折线图最为平缓稳定，直观的反映出本文改进算法在稳定性方面的优势。

4.2 实验二 脑部图像的分割

实验二同样利用上述方法对 7%高斯和 7%椒盐的混合噪声

的脑部图像分别处理，并对分割结果进行对比分析。图 3 是各算法对 7%高斯和 7%椒盐混合噪声的脑部图像的分割结果对比，图 4 是对个别图像的部分区域的细节放大对比，表 3 是不同算法处理相同噪声图像的耗时对比。实验软硬件环境及实验参数设置与实验一相同。

图 3b 是 7%高斯和 7%椒盐的混合噪声的脑部图像，图 3c、3d、3e 分别是 FCM 算法、自适应加权算法和本文改进算法处理后的结果。从图 3c 可以看到，FCM 虽然对细节保留最完整但是其去噪效果最差，图 3d、3e 均可较好的去除噪声，但图 3d 依然存有不错的分散噪声点，而图 3e 则极稀少。同时，从图 3a、3d、3e 三图中红色标记的细节处的连接对比可以看出，图 3a 图红框中图像是连接的，但图 3d 中连接断裂，图 3e 中

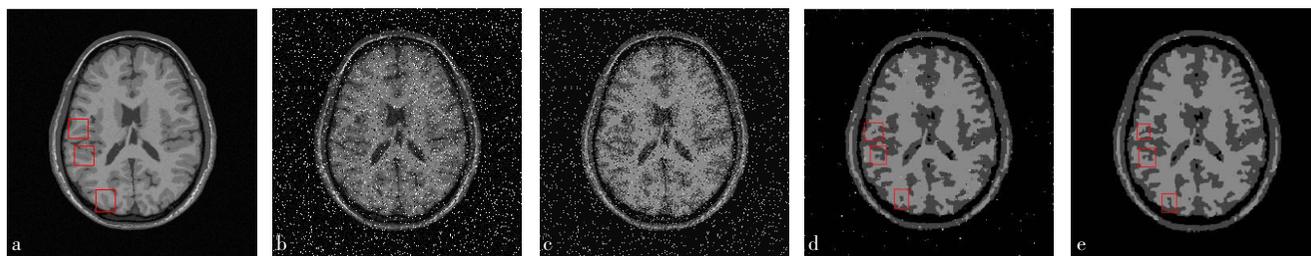


图 3 各算法对 7%高斯和 7%椒盐混合噪声的脑部图像进行分割的结果

(a: 脑部图像原图; b: 7%高斯+7%椒盐; c: FCM 处理后; d: 自适应加权处理后; e: 本文改进算法处理后)

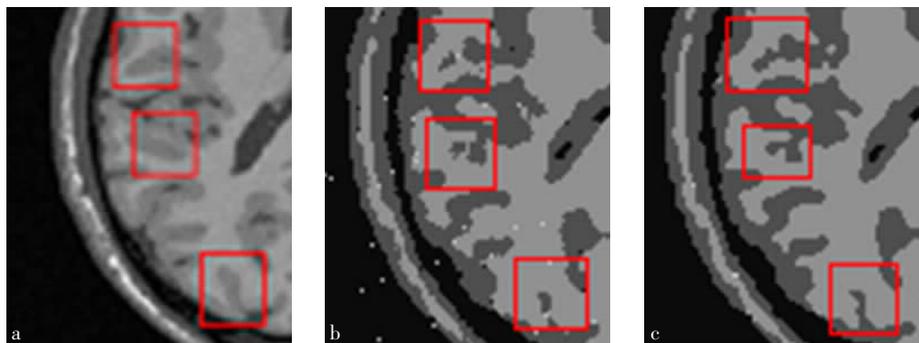


图 4 局部放大细节对比

(a: 脑部原图像; b: 自适应加权处理后; c: 本文改进算法处理后)

还保留了此连接。为更好说明这一点，将图 3a、图 3d、图 3e 三图部分区域放大对此，如图 4 所示。因此，本文改进算法在保留图像细节方面的效果也明显优于其它算法。

表 3 不同算法对相同噪声图像的处理耗时

项目	算法耗时 (s)	
	5%高斯合成图像	5%高斯脑部图像
FCM	3.19s	4.80s
自适应加权	5.44s	5.41s
本文改进	4.36s	4.98s

表 3 给出在 Intel Core2 DuoT5800 2.0G CPU、2G DDR2 内存环境下不同算法处理分别 5%高斯噪声的人工合成图像和脑部图像的消耗时间对比。

5 结论

本研究在传统 FCM 算法的基础上提出一种新的改进算法，该算法引入 Gibbs 随机场，将 Gibbs 随机场先验概率与隶属度的乘积作为改进后得图像像素的隶属度函数，在更好地平滑噪声的同时提高了对图像像素点的分类精确度。本研究通过对人工合成图像、脑部图像加入不同类型不同比例的噪声，对比传统 FCM 算法、自适应加权算法、本文改进算法的分割结果，清楚地表明了本文改进算法在去除噪声和精确分割方面的优越性，是一种简单高效的方法。

参考文献

- [1] Bezdek JC. Pattern recognition with fuzzy objective function algorithms [M]. New York: Plenum Press, 1981.
- [2] 丁震, 胡钟山, 杨静宇, 等. FCM 算法用于灰度图像分割的研究 [J]. 电子学报, 1997, 25 (5) : 39-43.
- [3] Xiang-Yang Wang, Juan Bu. A fast and robust image segmentation using FCM with spatial information [J]. Digital Signal Processing, 2010, 20 (4) : 1173-1182.
- [4] Stan Z.Li. Markov Random Filed Modeling in Image Analysis [M]. Tokyo: Springer, 2001.
- [5] J.M. Hammersley, P Clifford. Markov field on finite graphs and lattices [Z]. Unpublished Manuscript, 1971.
- [6] 程可嘉. 基于核函数的模糊聚类算法研究 [D]. 成都: 电子科技大学, 2009.
- [7] Derin H, Elliott H. Modeling and segmentation of noisy and textured images using gibbs random fields [J]. IEEE Trans Pattern Anal Mach Intell, 1987, 9 (1) : 39-55.
- [8] 冯衍秋, 陈武凡, 梁斌, 等. 基于 Gibbs 随机场与模糊 C 均值聚类的图像分割新算法 [J]. 电子学报, 2004, 32 (4) : 645-647.
- [9] 王顺风, 张建伟. 基于 Gibbs 场与模糊 C 均值聚类的脑 MR 图像分割 [J]. 计算机应用, 2008, 28 (7) : 1750-1752.
- [10] 林亚忠, 郝刚, 顾金库, 等. 结合随机场的自适应加权 FCM 改进方法 [J]. 计算机工程与应用, 2011, 47 (15) : 171-174.

(责任编辑: 常海庆)