

一种引入 ISODATA 动态聚类的医学图像中值滤波算法

席卫文袁 猛渊第一军医大学生物医学工程系袁广东 广州 510515冤

摘要 的 提出一种医学图像中值滤波算法袁改进经典中值滤波存在的不足遥方法 将模式识别中的 ISODATA 聚类引入到中值滤波算法中袁将分类的结果作为参数来确定如何进行中值滤波袁及是否需要进行中值滤波处理遥结果 本算法能够滤除严重的脉冲噪声干扰袁同时保持图像细节遥结论 本算法跟经典中值滤波算法相比袁提高了信噪比袁增强了图像质量遥

关键词 中值滤波 ISODATA 聚类 图像处理

中图分类号 R319;R814.3 文献标识码 B 文章编号 000-2588(2002)06-0558-03

A median filter algorithm importing ISODATA dynamic clustering for medical imaging

XI Wei-wen, ZHOU Meng

Department of Biomedical Engineering, First Military Medical University, Guangzhou 510515, China

Abstract: To improve conventional median filter algorithm employed in medical imaging, we proposed a new median filter algorithm importing ISODATA dynamic clustering for pattern recognition. The result of clustering was set as the parameters to decide if median filter was necessary and when it was, how the process was to be carried out. As shown in our test, this algorithm was capable of eliminating serious impulse noises and retaining thorough image details, therefore enhanced signal to noise ratio and quality of the images in contrast with the conventional median filter algorithm.

Key words: median filtering; ISODATA dynamic clustering; image processing

中值滤波是一种非线性信号处理方法袁971 年由 J.W.Jukey 首先提出并应用在一维信号处理技术中袁后来被二维图像信号处理技术所引用遥中值滤波对消除脉冲干扰以及图像扫描噪声最为有效袁且在实际应用中并不需要图像的统计特性袁算法简单袁所以近年来在图像处理领域应用广泛袁是图像处理技术中最常用的预处理技术之一遥但是中值滤波处理也在一定程度上模糊了图像边缘袁特别是在处理一些点尧线细节比较多的图像时遥本文提出了一种医学图像的自适应中值滤波算法袁引入模式识别中的 ISODATA 聚类袁根据分类结果来确定中值滤波的参数袁及是否需要进行中值滤波处理遥结果表明本算法能够滤除严重的脉冲噪声干扰袁同时保持图像细节遥

1 引入 ISODATA 聚类的中值滤波算法

1.1 中值滤波简介

中值滤波是用一个具有奇数点的滑动窗口 A 袁将窗口中心点的值用窗口所有点的中值代替遥例如有一维窗口各点的值分别为 $f_1=80$ 袁 $f_2=90$ 袁 $f_3=200$ 袁 $f_4=110$ 袁 $f_5=120$ 袁则窗口中心点的值 f_3 中值滤波后就应该是 120 袁用数学公式表示为院

$$f_3 = \text{Med}(f_1, f_2, f_3, f_4, f_5) \quad \text{渊冤}$$

通用的中值滤波表达式可记为院

$$f = \text{Med} (f) \quad \text{渊冤}$$

二维的中值滤波的窗口形状和尺寸对滤波效果影响较大袁常见的窗口形状有方形尧线形尧圆形尧十字形等袁窗口尺寸一般取 3~5 袁尺寸越大袁消除严重的脉冲噪声干扰作用越强袁但是越发导致了图像细节的丢失遥

1.2 ISODATA 聚类算法简介

ISODATA 聚类算法是一种动态聚类袁在模式识别领域里的应用比较广泛袁它是一种基于样本间相似性度量的非监督学习方法遥对同一类别的样本袁其特征向量应该是互相接近袁因此在不知道概率密度函数的情况下袁可以按照样本特征向量的相似性把样本集合聚类成几个子集袁这种划分的结果应该使某种表示聚类质量的准则函数达到一个极值遥 ISODATA 聚类算法的准则函数是误差平方和袁它将一个样本集聚类为 K 类袁若 N_i 是第 i 聚类 T_i 的样本个数袁 n_i 是聚类 T_i 中所有样本特征量的均值袁院

$$m_i = \frac{1}{N_i} \sum_{y \in T_i} y \quad \text{渊冤}$$

把 T_i 中的各样本特征值 y 与均值 m_i 间的误差平方和对所有类相加后为院

$$J_k = \sum_{i=1}^k \sum_{y \in T_i} ||y - m_i|| \quad \text{渊冤}$$

对于同一样本集合袁不同的聚类袁 的值自然不同袁 J_k 极小的聚类就是误差平方和准则下的最优结果遥

1.3 引入 ISODATA 聚类的中值滤波算法

假设 y_{ij} 代表噪声图像中图像点 (i, j) 的灰度值以它为中心的窗口 S 是一个 3×3 的方形窗口利用 ISODATA 算法根据灰度值这一特征量将这一方形窗口的像素点动态分为 K 类 ($K < 4$) 其步骤如下

首先要对样本集进行初始分类先选择窗口 S 中灰度值最大和最小的两点分别作为两类聚类 T_1 和 T_2 的核心

$$m_1 = \text{Max}_{y \in S}(y), m_2 = \text{Min}_{y \in S}(y), K=2$$

把窗口 S 中所有的点按照最小平方误差准则分到 T_1 和 T_2 这 K 个聚类中

$$e_i = (y - m_i)^2, i=1, 2, \dots, K$$

如果 $e_j = \text{Min}(e_i)$, 则 $(i, j) \in T_j$

更新各聚类的灰度均值

$$m_i = \frac{1}{N_i} \sum_{y \in T_i} y, i=1, 2, \dots, K$$

其中 N_i 为第 i 个聚类 T_i 中像素点的个数

对每一个聚类 T_i 求出它的标准偏差

$$s_i = \sqrt{\frac{1}{N_i} \sum_{y \in T_i} (y - m_i)^2}$$

比较标准偏差 s_i 与标准偏差阈值

标准偏差阈值的取值可以自定义令 $\theta = (\text{Max}(y) - \text{Min}(y))/2$ 如果 $s_i > \theta$ 则把聚类 T_i 分裂为两个聚类 T_i^+ 和 T_i^- 它们的聚类中心点的灰度值 m_i^+ 和 m_i^- 分别为

$$m_i^+ = m_i + a * s_i, m_i^- = m_i - a * s_i (0 < a < 1)$$

按照步骤 2 所示的方法使用最小平方误差准则使得原来聚类 T_i 的像素点分别重新聚类到 T_i^+ 和 T_i^- 且令 $K=K+1$

对于所有的聚类计算它们两两之间的距离

$$d_{ij} = (m_i - m_j)^2, i, j=1, 2, \dots, K-1; j=i+1, \dots, K$$

比较 d_{ij} 与合并参数阈值

合并参数阈值的取值可以自定义令 $\lambda = (\text{Max}(y) - \text{Min}(y))/6$ 如果 $d_{ij} < \lambda$ 则将聚类 T_i 和 T_j 合并成一个聚类 T_k 新的聚类的灰度均值为

$$m_k = \frac{1}{N_i + N_j} (N_i m_i + N_j m_j), K=K-1$$

重复步骤 1 以下的操作迭代 L 次 一般大于 5 就能够得到较好的聚类结果 以图像点 (i, j) 为中心的窗口 S 经过 ISODATA 聚类后设定中心点所在的聚类记为 T_k 其所包含的像素点的个数记为 n 中心点的灰度值为 x 对这一方形窗口进行中值滤波处理分为以下几种

$$K=1, n=9$$

整个方形窗口的灰度值分布比较均匀中心点的灰度值保持不变

$$K=2, n=1$$

T_1 只包含邻域中心这一点中心点很可能就是灰度均匀分布的邻域中孤立的噪声点对其余的像素点所在的聚类 T_2 进行中值滤波所得的灰度值作为图像点 (i, j) 的灰度值 x

$$x = \text{Med}_{y \in S}(y)$$

$$K=2, 3 < n < 6$$

邻域窗口中聚类后两类的个数相差不大由于相邻像素之间有一定的空间相关性所以邻域窗口中存在边缘的可能性很大为保持边缘细节不应该对整个方形窗口进行中值滤波 为达到锐化的效果还可以加入一些锐化处理比如

$$x = x + (m_1 - m_2) * 0.5$$

$$K=2, n=8$$

中心点位于灰度值分布比较均匀的聚类 T_1 中则对聚类 T_1 进行中值滤波

$$x = \text{Med}_{y \in S}(y)$$

$$K=3, n=1$$

中心点很可能是孤立的噪声点对整个方形窗口的像素灰度值进行中值滤波

$$x = \text{Med}_{y \in S}(y)$$

$$K=3, n < 6$$

分两种情况

如果另外有一个聚类的样本个数也是 3 到 6 之间处理同

保持中心点的灰度值不变

$$K=3, n=7$$

处理同

其它

保持中心点的灰度值不变

2 实验结果

采用的原始图像是深圳安科公司提供的 256 级灰度的脑部 MRI 图 1 为加入了随机噪声的图像 在 PIII333 56MSDRAM 0G 硬盘的微机上用 VISUALC++6.0 编程分别对图 2 进行经典中值滤波和本文提出的引入 ISODATA 聚类的中值滤波处理得到图 3 和图 4 图 3 虽然很好地消除了孤立噪声的污染但是造成了一定的图像细节损失 图像变得模糊 图 4 则既消除了脉冲噪声带来的干扰又保持了图像的边缘细节甚至还增强了图像的边缘细节

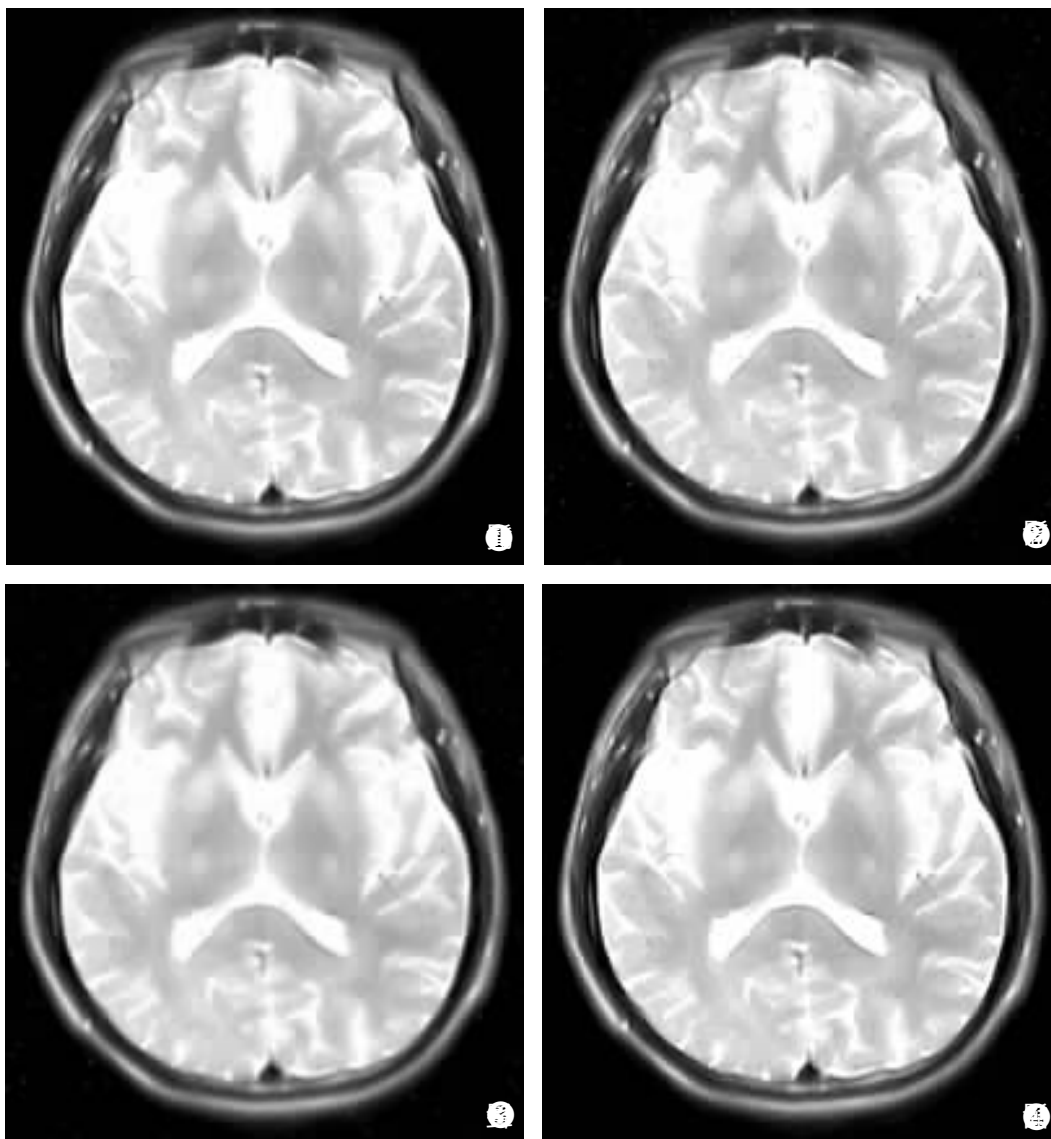


图 1 原始图 Fig.1 Original image

图 2 加入随机噪声的图像 Fig.2 Images blurred with random noises

图 3 经典中值滤波的图像 Fig.3 Result of conventional median filter for the blurred image

图 4 改进中值滤波的图像 Fig.4 Result of improved median filter for the image

3 结论

经典的中值滤波器只能一致地应用到整个图像,无论图像点的好坏均受到影响,这样必将会破坏许多好的图像细节。本文提出了一种引入 ISODATA 聚类的中值滤波算法,它首先对邻域窗口的图像点进行聚类,然后根据聚类情况对窗口内是否存在边缘,是否存在孤立噪声点进行分析,根据分析结果进行全局或者局部的中值滤波,或者不进行中值滤波。实验结果表明,该方法在抑制图像严重的脉冲噪声干扰和细节特征保持方面有较好的性能,与经典的中值滤波相比较,本算法由于增加了聚类分析的过程,处理过程在

时间上的开销要大一些,但是在提高信噪比,增强图像可懂性方面有了明显改进,并且算法简单,易于实现,在医学图像处理领域有实际应用价值。

参考文献

咱暂 赵荣椿. 数字图像处理导论[M]西安:西北工业大学出版社, 1995.130-80.
 咱暂 DonohoDL. De-noising by soft-thresholding [J] IEEE Trans Info, 1995, 41(4):613-27.
 咱暂 边肇祺, 张学工. 模式识别[M]北京:清华大学出版社, 1999.230-50.
 咱暂 Dutt V. Adaptive speckle reduction filter for log-compressed B-scan images [J] IEEE Trans Med Imaging, 1996, 15(6):802-13.