

文章编号: 1005-0523(2009)06-0065-04

基于均值聚类的体绘制中传递函数的自动产生

马得安, 蒋先刚, 陶龙凤

(华东交通大学 基础科学学院, 江西 南昌 330013)

摘要:采用均值聚类的算法为三维数据场直接体绘制的传递函数中的各类分界自动产生关键点。该方法为传递函数关键点的初选提供了合理的依据和参考,将器官组织的灰度自动分为 c 类,传递函数的关键点取在边界点处,并使分界点处的透明度为较大值,这样使器官组织的分割自动达到较好状态。实验显示,通过自动产生的传递函数比手动选择的传递函数提供了表达更清晰组织的重构的显示效果。

关键词:K-均值算法;传递函数;曲线控制点
中图分类号:TP301 **文献标识码:**A

体绘制是 1 种重要的三维数据场可视化算法,它是 1 种通过计算彩色半透明体的二维映射来呈现三维空间样条函数的技术。体绘制算法的关键是传递函数的设计,它将三维数据场的数据值转换为光学成像参数,决定了投影图像的质量。

目前传递函数设计方法主要有手动调节法、图像中心法、数据中心法和对象中心法这 4 种方法。其它的有采用 ISODATA 聚类算法对数据场进行聚类,还有利用神经网络和支持向量机等机器学习算法对数据场分类,并设计成高维传递函数。传递函数的自动产生将以图片内部表达组织的类型数、整体分布、局部分布和梯度分布情况通过计算得到,本论文采用均值聚类算法自动产生传递函数曲线的关键点而进行三维数据场直接体绘制,聚类算法以重构切片的图像灰度分布特性数据为分析对象。

1 三维数据场中的均值聚类传递函数的作用

在直接体绘制中,传递函数的作用是将三维数据场的数据值转换为光学成像参数,建立起数据点与光学特征的映射关系。数学上,传递函数可以定义为三维数据场的数据属性到光学属性的映射:

$$\tau: D_1 \times D_2 \times \cdots \times D_l \rightarrow O_1 \times O_2 \times \cdots \times O_m$$

其中: $D_i (i = 1, 2, \cdots, l)$ 是传递函数的定义域,表示三维数据场的数据属性,数据属性是三维数据场自身的数值特征,它可以是采样点的数据值,如 CT 图像的灰度值,也可以是局部采样点数值计算的结果,如梯度幅值、二阶方向导数、曲率等; $O_i (i = 1, 2, \cdots, m)$ 是传递函数的值域,表示进行可视化的光学属性如颜色 (RGB)、不透明度 (alpha)、阴影参数 Phong、反射率、折射率等; τ 表示数据属性转换为光学属性的映射规则。设计传递函数就是根据可视化的需求,选择和设计合适的数据属性和光学属性,建立它们之间的映射关系^[1]。

K 均值聚类算法对数据场进行分类,采样点的任何数据属性都可作为算法的输入。它根据输入数据空间的距离对数据场进行分类,输出 1 组具有类似特征的采样点簇和簇均值。为了降低计算量,可先从数据场中任意地选取 1 组采样点进行预分类得到簇均值,然后再对整个数据场分类。被选择的采样点必须包含整个数据场的直方图的典型分布,采样点数目由数据场的大小确定,K 均值聚类算法将传递函数的定义域拓展到了簇空间,程序中的用户界面还允许用户对分类的结果进行修改和调节。

收稿日期: 2009-06-12

基金项目: 江西省教育厅科学技术项目 (GJJ09212)

作者简介: 马得安 (1987-), 男, 山东日照人, 硕士研究生, 研究方向为图像处理与信息技术。

1.1 K-means 算法

K-means 算法的思想就是使得被划分到同一簇的对象之间相似度最大,而不同簇之间的相似度最小。

K-均值算法^[2]的一般步骤如下

(1) 假设给出 n 个数据点 $U = \{u_1, u_2, \dots, u_n\}$, 其中 $u_j = (x_{j1}, x_{j2}, \dots, x_{jm}) \in R^m$. 取定分类数 $c (2 \leq c \leq n)$, 并初始化 $A^{(0)} \in M_c$, 其中

$$M_c = \{A \mid A = a_{ij} \in V_{c \times n}, a_{ij} \in \{0, 1\}, \sum_{i=1}^c a_{ij} = 1, 0 < \sum_{j=1}^n a_{ij} < n\}$$

(2) 当迭代次数为 $l (l = 0, 1, 2, \dots)$ 时, 计算聚类中心向量

$$v_i^l = \frac{\sum_{j=1}^n a_{ij}^l u_j}{\sum_{j=1}^n a_{ij}^l}, \text{ 其中 } a_{ij}^l = A^l, i = 1, 2, \dots, c$$

(3) 用下式将 $A^l = a_{ij}^l$ 更新为 $A^{(l+1)} = a_{ij}^{(l+1)}$

$$a_{ij}^{(l+1)} = \begin{cases} 1 & \|u_j - v_i^l\| = \min_{1 \leq i \leq c} (\|u_j - v_i^l\|) \\ 0 & \text{其它} \end{cases}$$

(4) 比较 A^l 和 $A^{(l+1)}$, 若 $\|A^{(l+1)} - A^l\| < \epsilon$, (ϵ 是一个非常小的常数), 则停止算法; 否则, 令 $l = l + 1$, 返回(2)。

1.2 器官组织分割的边界点的获取及传递函数关键点的取值

在三维数据场中, 通常假设同 1 种组织具有相似的灰度值, 不同的组织具有不同的灰度值, 则不同组织之间的边界则是 1 个关键问题, 通过 K-均值聚类算法对图像颜色进行分类后, 不同的组织之间已具有了不同的灰度, 我们将通过在不同组织边界间的梯度分布来进行处理, 那么在组织的边界处灰度值将发生显著的变化, 这样就达到分割开不同器官组织的目的。如图 1 所示, 切片的灰度直方图分布如曲线显示, 它表示不同器官组织和体液的一种分布情况, 灰度值变化在 0 ~ 255 之间, 通过均值聚类的颜色分类可以分为 4 类组织, 它们分别在灰度区域 1-2, 2-3, 3-4, 4-5 区间, 4 类的均值聚类的中心分别为 A, B, C 和 D, 这样, 为了表达每相连的两类组织的区别而选择透明度传递函数多边形折线的谷底为每相连的两类组织的分界点 2, 3, 4, 通过计算机自动产生的透明度传递函数多边形折线如图中粗线所示, 这样传递函数的关键点的取值为各分类的边界点, 也就是图 1 中的谷底点, 在自动产生的传递函数的多边形折线的基础上经过手动细调折线上的各点并与亮度传递函数和彩色传递函数的调节相结合就可产生清晰和组织分明的三维物体图像。

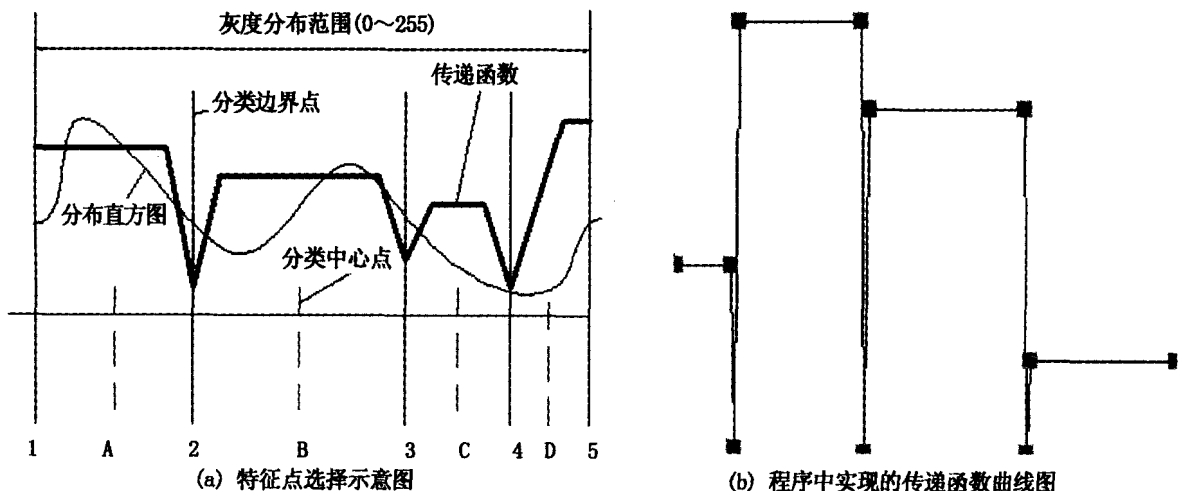


图 1 通过均值聚类自动产生的传递函数折线关键点

2 实验结果对比

本研究采用 Delphi 7 来实现算法^[3]。实验数据为 $512 \times 512 \times 128$ 的骨头 CT 扫描数据。为了提高分析速度,可以随机地选取部分切片数据进行分析,当然选取的切片数据必须代表器官组织的整体分布情况,本文随机抽取了 50 组数据进行检验。首先将三维数据灰度化,将其映射到由灰度范围 $0 \sim 255$ 的特征空间中,然后采用均值聚类算法在特征空间中对采样点进行聚类。

为了更好地验证算法的效果,图 2 给出了数据场中某一切片的用均值聚类传递函数自动实现的结果图。图 2(a)为胸部原始切片,图 2(b)为手动调节传递函数作用后的分割结果,图 2(c)为自动聚类后传递函数作用的分割结果。切片数据聚为 4 类:骨头、肌肉组织、骨髓和背景。图 2(c)相对于图 2(b)分类更好,而且层次效果也更加明显,产生的传递函数曲线的关键点为各器官组织的分割提供初步方案和关键参数,它是由均值聚类方法自动产生,该方法减少了医生手工调节中传递函数曲线的不确定性,同时提高了工作效率。

图 3 为一胸部三维重构显示效果的比较效果图,图 3(a)为胸部切片依 45° 方向线排布的传递函数作用的重构效果图,图 3(b)通过聚类后自动产生传递函数作用的重构效果图,在放大区域可以明显看出经过均值聚类传递函数作用后的图片的器官组织纹理层次分类的效果是很明显的,而且血管与原始图片相比也更加清晰,出现断裂的也有修复的效果。

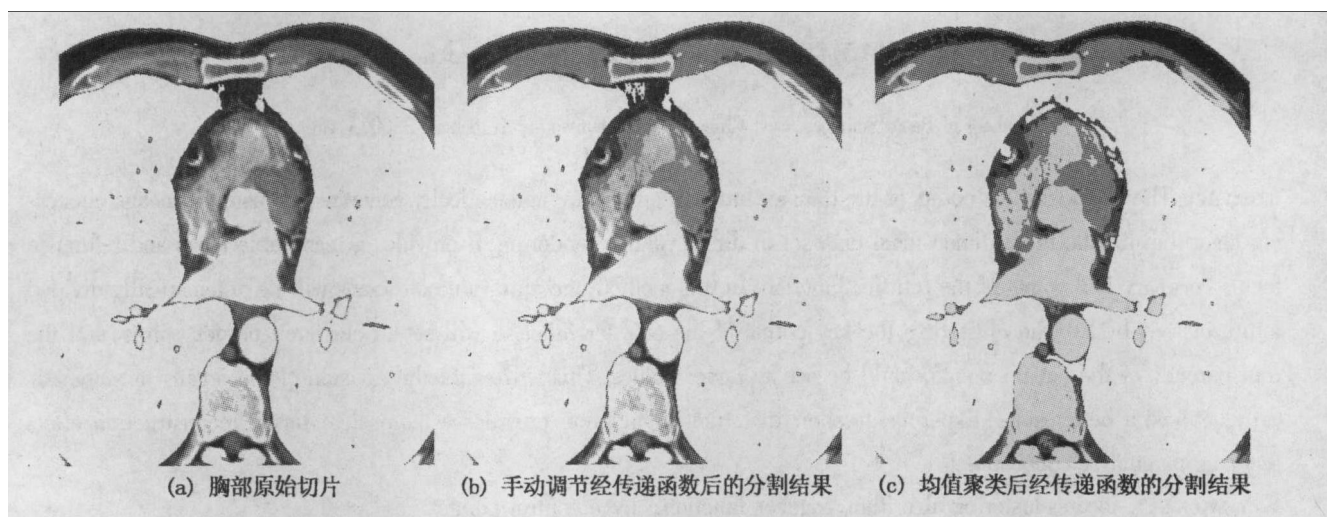


图 2 胸部切片经传递函数后的比较效果图

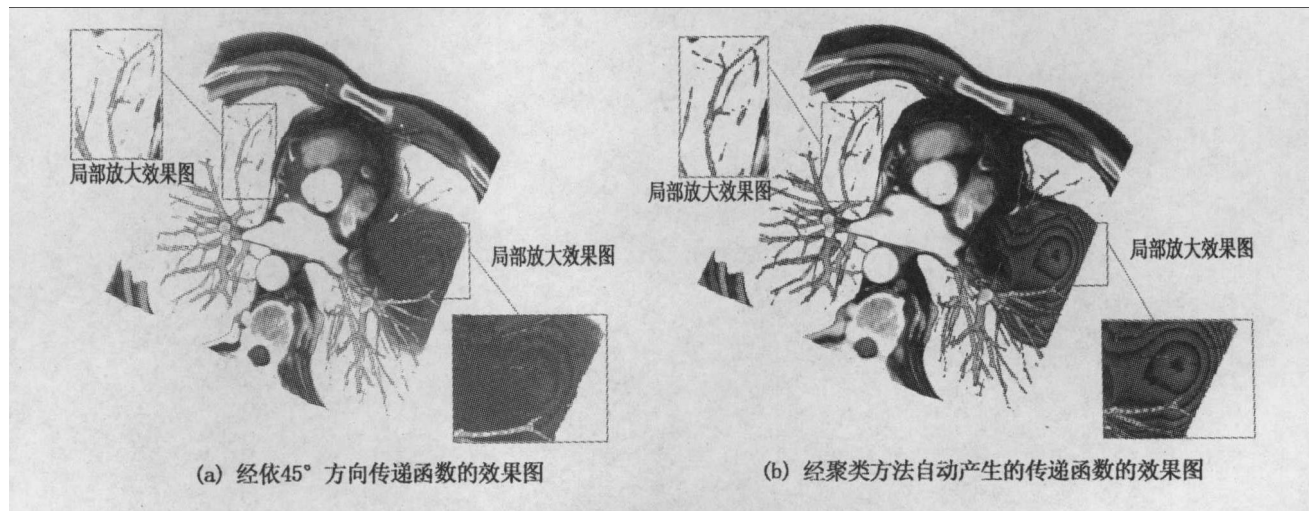


图 3 胸部组织三维数据场重构显示效果图比较

3 结论

利用均值聚类算法提出了1种三维数据场中通过自动聚类产生传递函数的方法,该方法通过切片的灰度分布而自动将器官组织分为 c 类,且在分类边界通过调节较大透明度,使层次更加清晰而达到较好三维显示效果。均值聚类对三维数据灰度域的合理分界为传递函数关键点值的划定提供了粗划分界限。在此传递函数曲线的轮廓上各点的作用下,三维数据场将以更加清晰的组织分割效果呈现出来,为经验性传递函数的产生提供初步的关键点选择,提高了三维重构软件在诊断时的临床应用性。

参考文献:

- [1] 周芳芳,樊晓平,叶 榛.基于自适应宽带均值漂移聚类算法设计传递函数[J].信息与控制,2007,36(5):585-591.
- [2] 蒋先刚.数字图像模式识别工程软件设计[M].北京:中国水利水电出版社,2008.
- [3] 刘金春,蒋先刚,詹雪峰.均值聚类在三维重构图像预处理中的应用[J].微计算机信息,2007,24(5-3):303-305.

Automatic Generating of Transfer Function Based on K-means Clustering in Direct Volume Rendering

MA De-an, JIANG Xian-gang, TAO Long-feng

(School of Basic Sciences, East China Jiaotong University, Nanchang 330013, China)

Abstract: The boundary key points of the transfer function curves are automatically generated by using k-means clustering algorithm for the three-dimensional data set in direct volume rendering. It provides a reasonable basis and reference for the primary key points of the transfer function. In this method, the gray value of organs will be automatically divided into c classes by k-mean clustering, the key points of the transfer function will set in clustered border points, and the transparency of the cut-off points would be set up larger values. This will enable the division of the organs to automatically achieve a better state. Experiments show that transfer function provides a more clear tissue reconstruction effect than the manual selection does.

Key words: K-mean clustering algorithm; transfer function; curve control points

(责任编辑:刘棉玲)