

PKU Visualization Blog

北京大学可视化与可视分析博客

一种基于三维卷积稀疏编码的体渲染智能系统方案 (An Intelligent System Approach for Probabilistic Volume Rendering using Hierarchical 3D Convolutional Sparse Coding)

作者: Fan Hong 日期: 2017年12月26日

直接体渲染是一种表现三维体数据的强大可视化手段。过去数十年间, 研究者们进行了大量的研究, 来提高体渲染的速度与质量。其中, 传递函数的设计是影响这两个指标的一个重要因素。传递函数将体素的值映射到诸如颜色、不透明度、可见性等光学性质。然而, 要设计一个好的传递函数通常面临着很大的困难, 对非专业使用者更是如此。绝大多数传递函数设计的方法都依赖用户大量的交互, 用户需要根据数据的统计特征, 通过不断试错, 以求达到好的效果。这里的统计特征通常是数据某些属性的直方图。在本文中, 作者提出了一种基于机器学习方法: 三维卷积稀疏编码, 能对体素进行精确分类, 从而得到优秀的体渲染结果。同时, 基于智能系统的交互方式, 能让用户通过直观的输入, 来完成传递函数设计过程。

近年来, 基于多维度直方图的传递函数因为其卓越的特征空间分类能力, 变得越来越受用户喜爱。但是, 由于传递函数空间与数据空间之间越来越大的感知差距, 也让操作多维传递函数空间变得十分复杂。本文一方面采用基于智能系统的方法, 让用户通过笔刷工具输入少量信息, 来完成传递函数设计过程, 减少了用户的交互负担。另一方面, 采用机器学习算法, 自动从体数据中提取更高维度的特征, 能更精确地对体素进行分类, 并能更好地对抗数据中的噪声。

首先, 我们来看一下什么是卷积稀疏编码。卷积稀疏编码是字典学习的一种, 其目的是希望用一组卷积过滤器 (convolutional filters) 来近似一张图像 s : $s = \sum_k d_k * x_k$ 。其中的 d_k 就是卷积过滤器, 也称作字典中的一个原子 (atom), 而 x_k 则是与卷积过滤器对应的稀疏图 (sparse map), $d_k * x_k$ 被称作响应图 (response map)。所以图像 s 就由响应图之和来近似了。 x_k 之所以被称之为“稀疏”, 则是因为算法要求 x_k 里各分量的数值尽量小, 具体求解中用1-范数来约束, 即所有分量绝对值之和。在以上两个要求之下, 求解卷积稀疏编码可以转化为一个最优化问题, 其能量函数如下:

$$\min_{d,x} \frac{\alpha}{2} \|s - \sum_k d_k * x_k\|_2^2 + \lambda \sum_k \|x_k\|_1, s.t. \|d_k\|_2^2 \leq 1,$$

其中第一项要求响应图之和尽量接近图像, 第二项则是对 x_k 的稀疏正则化项, 限制条件则是要求每个原子 d_k 的长度在单位长度以下。求解卷积稀疏表示的过程, 请参照文中所引用的论文。卷积稀疏编码既可以对单张图像进行求解, 也可以对一组图像进行求解, 只需要将能量函数进行相应修正即可。图1展示了对一组图像求解出来的包含64原子的字典。从中可以看到, 大部分原子都有着条纹状的形式, 它们能匹配图像中的边。

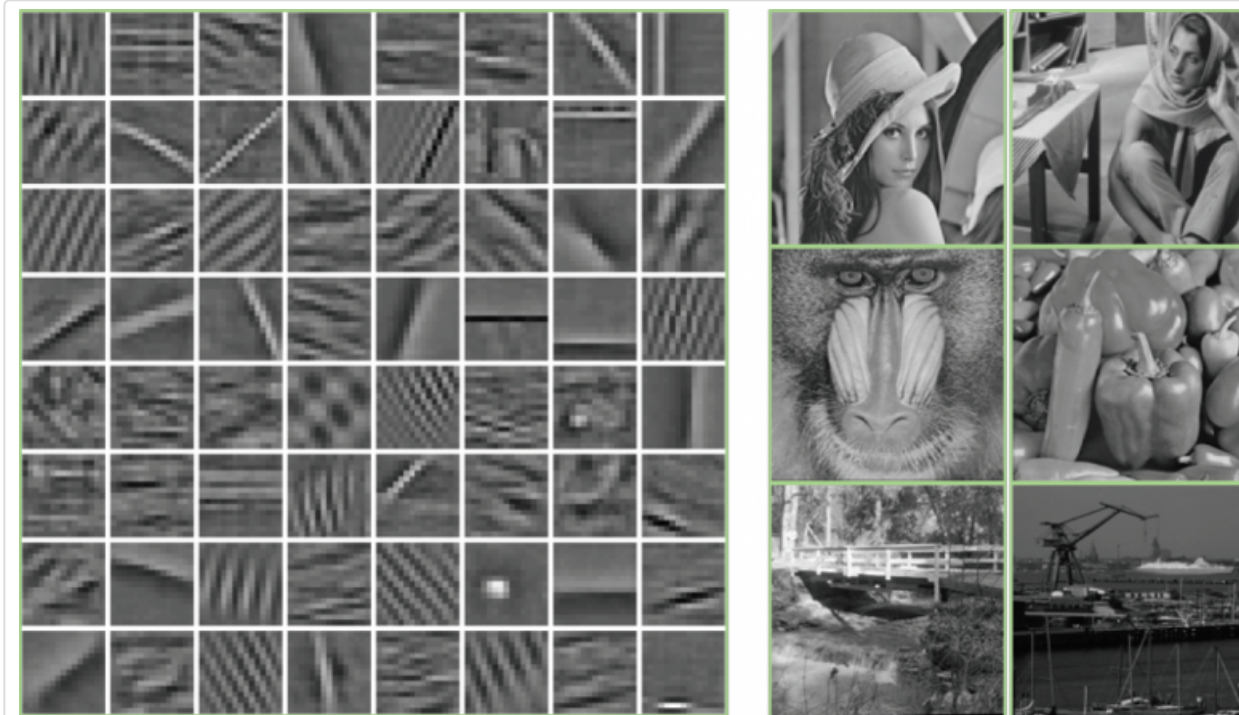
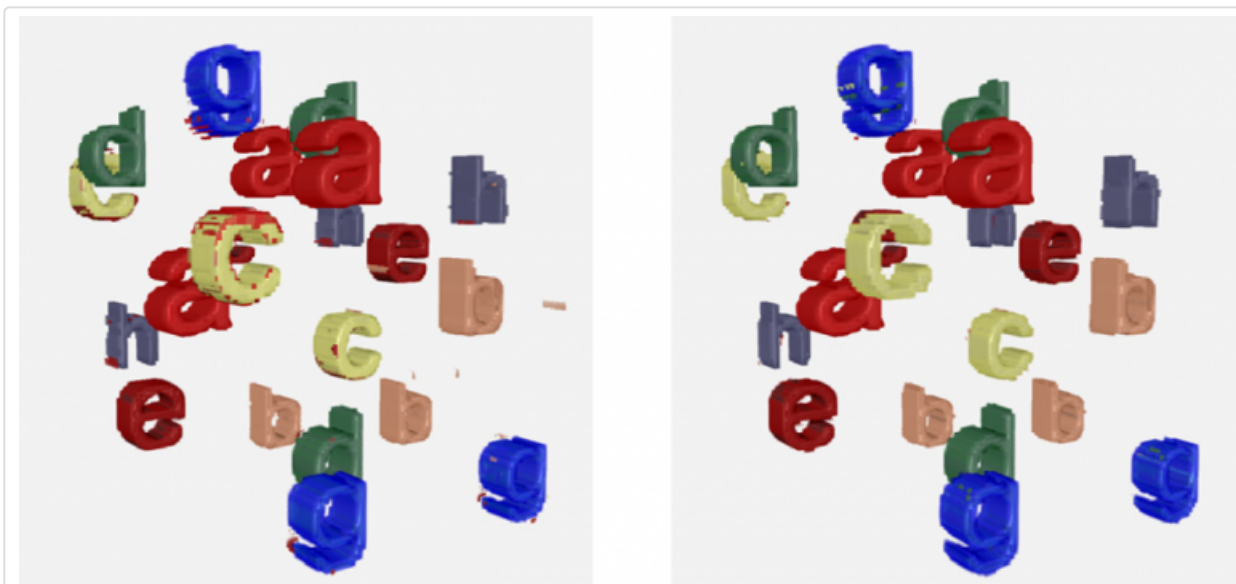


图1: 从一个图片集中学习到的一个包含有64个原子的字典

将卷积稀疏编码拓展到三维体数据并不困难。三维卷积稀疏编码的能量方程如下:

$$\min_{d,x,g,y} \frac{\alpha}{2} \|v - \sum d * x\|_2^2 + \lambda \sum \|y\|_1, \text{ s.t. } \|g\|_2^2 \leq 1, x - y = 0, g = Proj(d).$$

其中, 与二维情况相比, 我们的原子和稀疏图都变成了三维, 对原子的长度限制也修正为对其二维投影 g 的限制。求解方法这里同样不赘述, 请参考文中引用的论文。为了避免求解过程陷入局部最小值, 作者还引入了“blackout”策略。即在迭代求解过程中, 随机将一部分原子重置为0, 或者将一部分能量最大的原子重置。在图2中能看到, 加入blackout策略后, 根据字典进行体素分类的结果精度更高。



(a) No regularizer

(b) With blackout

图2: 体素分类准确率的比较 (a)没有正则化 (b)使用blackout正则化

作者接着将三维卷积稀疏编码拓展到多尺度上。作者发现当使用不同大小的卷积过滤器时, 它们学习到的特征也有着明显的不同, 如图3所示。并且, 基于多尺度卷积过滤器进行体素分类的效果要比单尺度的要更好, 如图4所示。

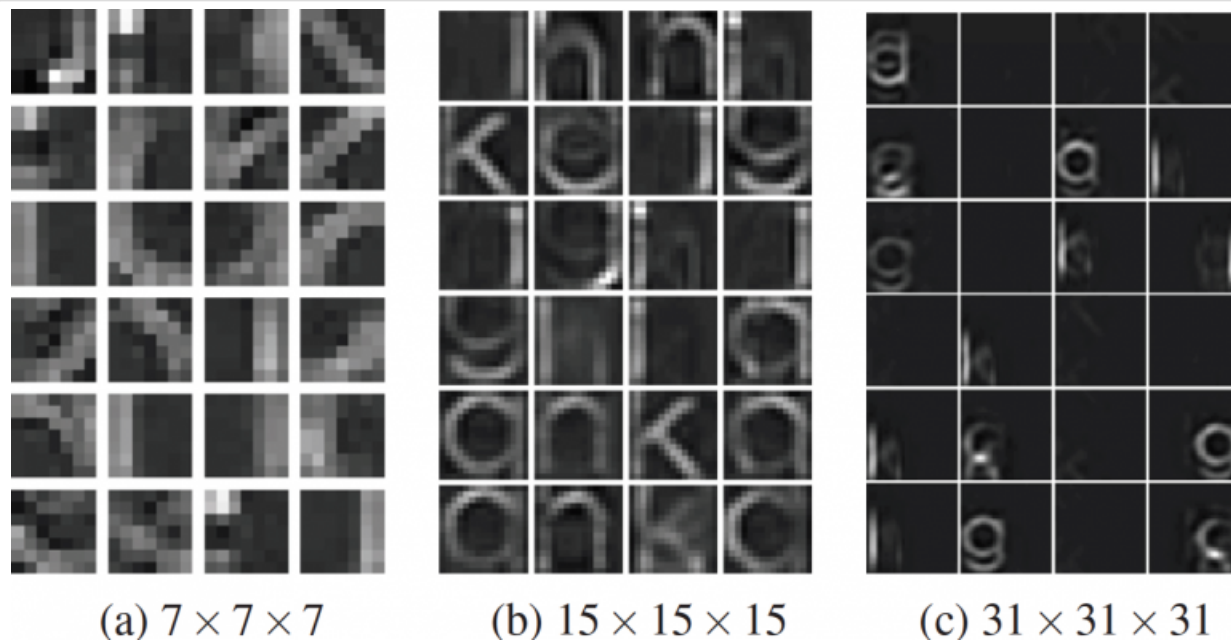


图3: 从alphabet数据集中学习到的多尺度的字典

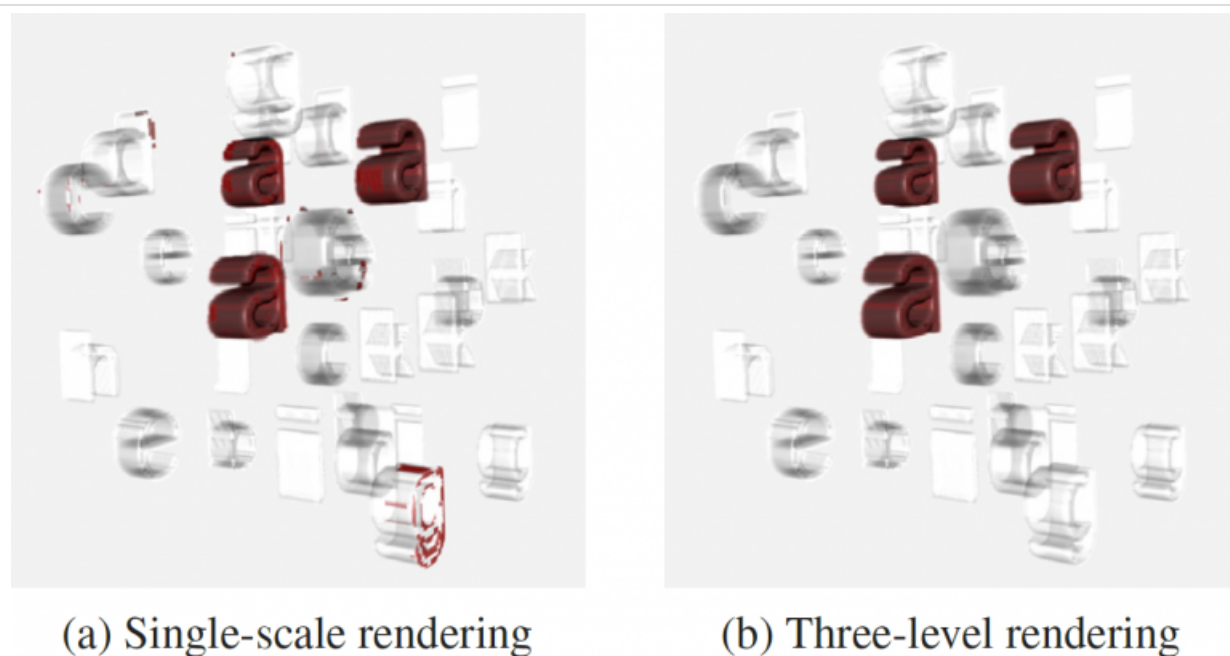


图4: 使用(a)单尺度字典 (b)多尺度字典进行体素分类后, 体渲染结果的比较

但作者在实现中并不是直接采用多尺度的卷积稀疏编码, 而是提出了一种层次式的做法, 如图5所示。首先用较小的卷积过滤器 ($7 \times 7 \times 7$) 来近似原体数据 v , 得到一组原子字典和稀疏图, 以及近似结果 v_1 。再用中等大小的卷积过滤器 ($15 \times 15 \times 15$) 来近似上一轮中的近似体数据 v_1 , 得到一组原子字典和稀疏图, 以及近似结果 v_2 。最后用较大的卷积过滤器 ($31 \times 31 \times 31$) 来近似 v_2 , 得到一组原子字典和稀疏图。

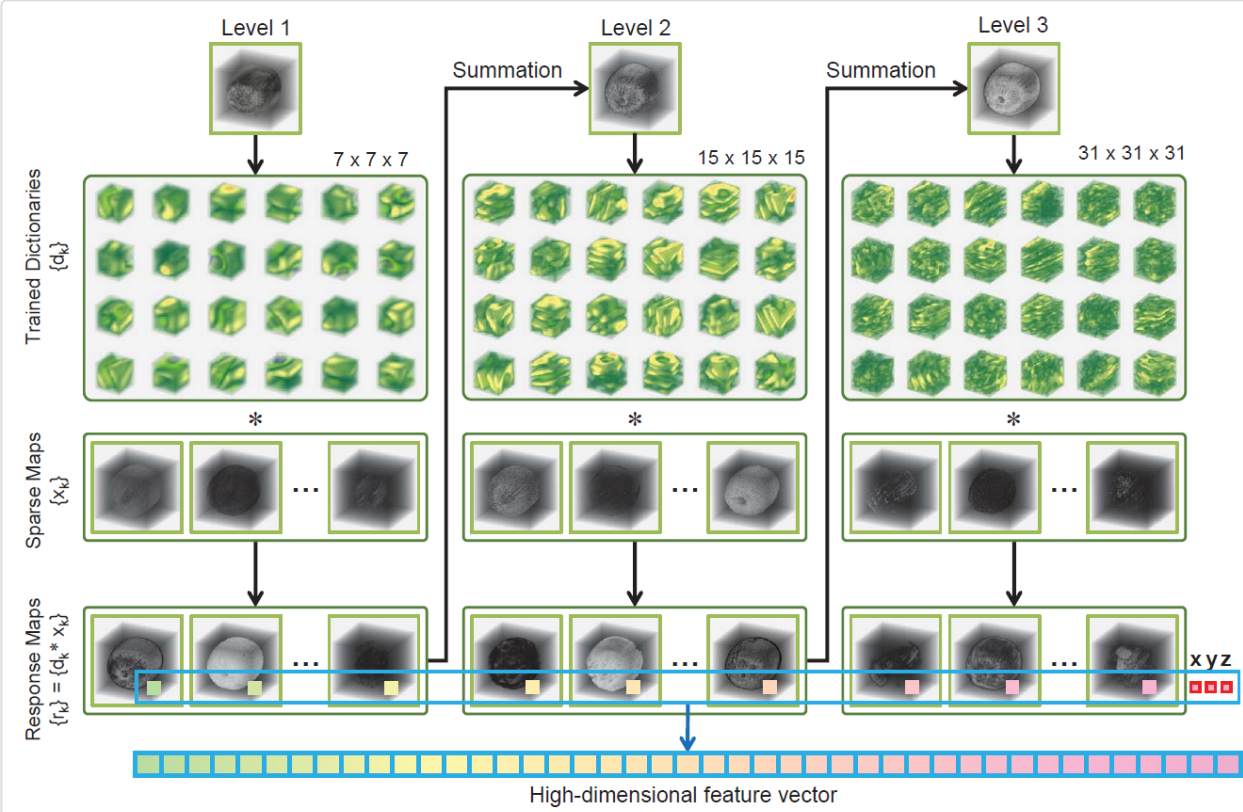


图5：层次卷积稀疏表示以及高维特征向量的构造

这里作者限制字典中原子的数目为24，相应地稀疏图数目也为24。这样，对于体数据中的每一个体素，其在三个层次的各24个稀疏图中各有一个取值，它们一起构成了一个72维特征向量。再加上每个体素的三维坐标，就构成了一个75维的特征向量。基于这个高维特征向量，作者实现了一个智能系统来完成体素的分类与传递函数设计任务。

用户在图6所示的界面中，用笔刷工具在体数据各个切面上选择体素进行分类，并给每个类别按属于这个类别的概率设计一维传递函数。然后，作者采用随机森林算法，以用户标定的体素为输入，根据75维特征向量对所有其他体素进行分类，得到各个体素属于各个类别的概率。然后，各个体素根据其属于各个类别的概率以及各个类别的一维传递函数，加权得到最终的颜色。最后采用体绘制算法得到体数据渲染结果。

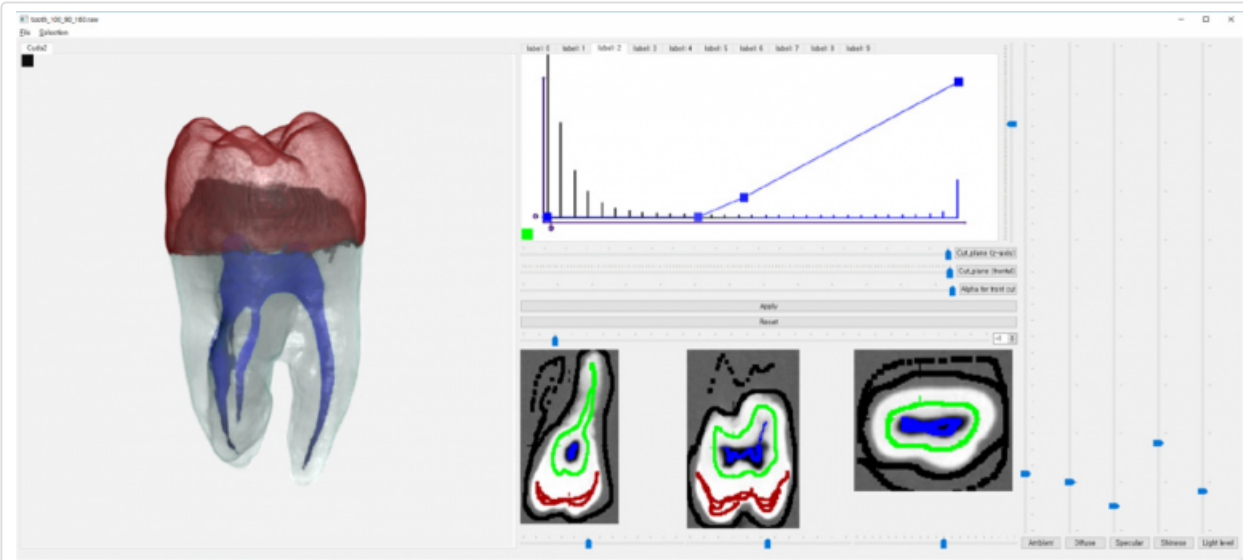
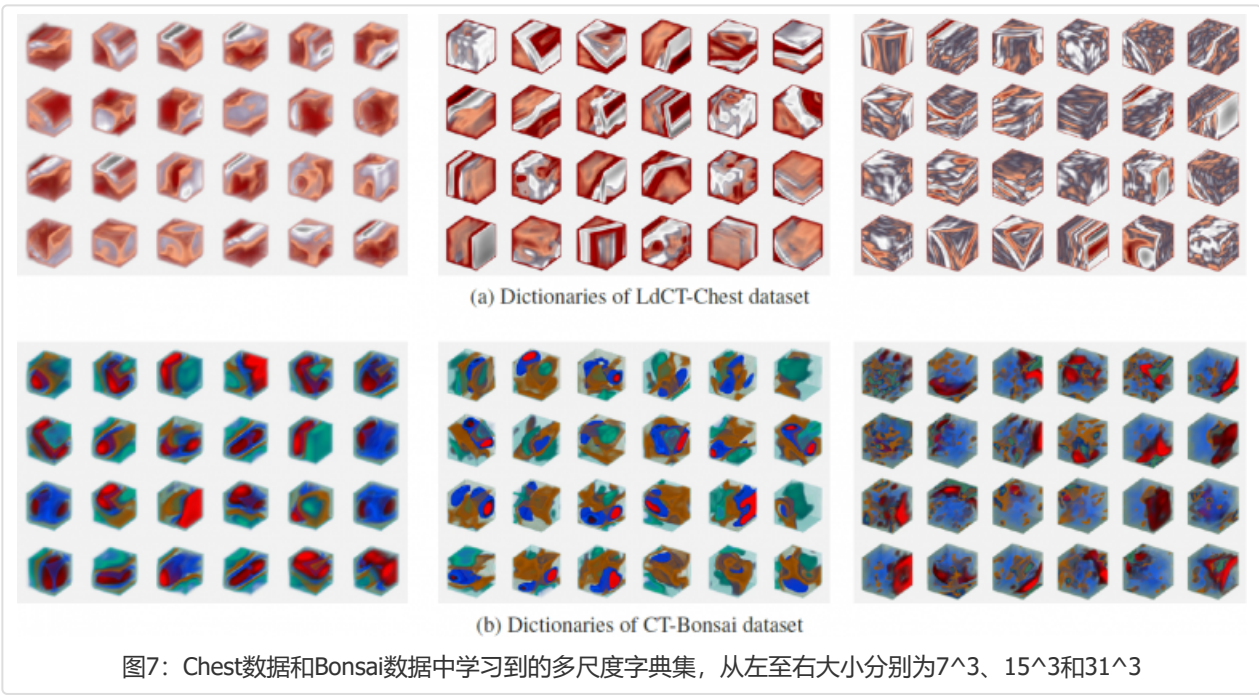


图6：本文智能系统的用户界面

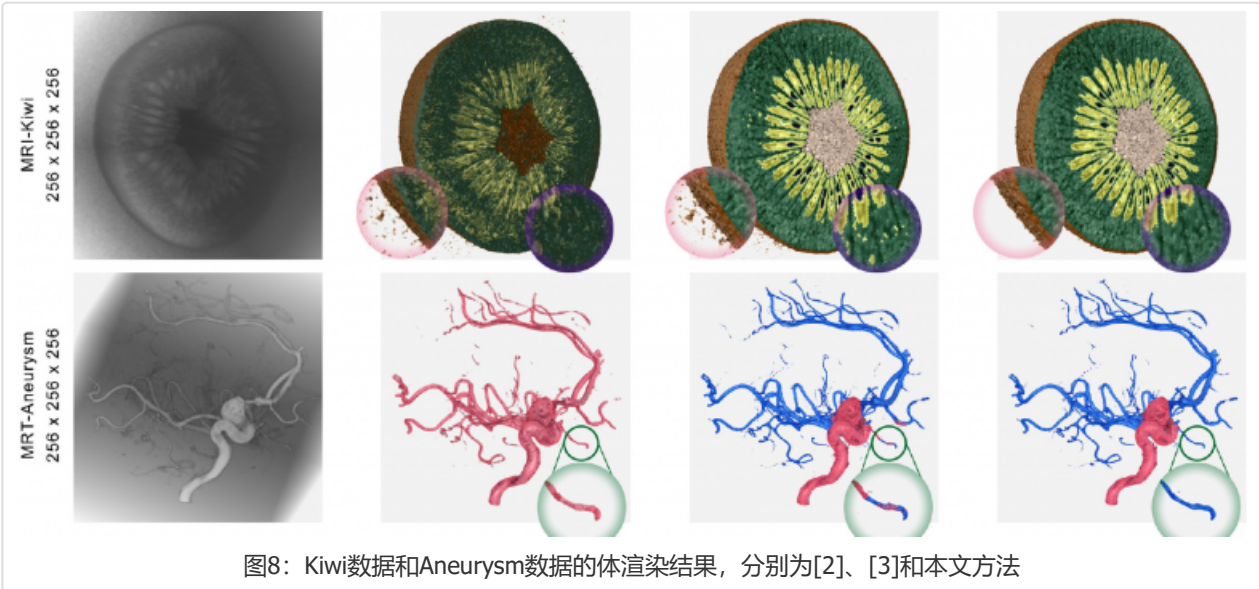
在结果中，作者首先对学习到的三维字典进行可视化，如图7所示。其中，在Chest数据中，能看到在精度较高时，学习到的原子能对应到骨头及其周围的肌肉组织。其他数据也能观察到类似的结果。



在体渲染的结果比较中, 作者首先测试了一个带噪声的合成数据, 结果如图8所示。已有的两个方法[2,3]对噪声均没有特别鲁棒, 而本文的方法能完全避免噪声。



对于真实体数据的渲染结果, 如图9所示。能看到基于3维传递函数设计的方法[2], 数据中的某些结构特征都无法完全识别出来, 例如Kiwi数据中的种子部分。而基于11维传递函数的方法[3], 则在某些体素中无法完全正确分类。本文的方法则都取得了很好的效果。更多数据的测试结果可以参看论文原文。



最后，作者还通过用户研究的形式，通过用户的实际操作，验证了本文系统的易用性以及准确性。

总起来看，本文提出的卷积稀疏编码与智能系统，将传递函数的直接设计问题转变为一定用户输入下的高维特征向量分类问题，既提高了传递函数的效果和对噪声的鲁棒性，又在系统的易用性上有了很大的进步。尽管方法本来与深度学习没有直接联系，但其中的层次卷积稀疏编码在一定程度上模仿了深度卷积神经网络中行为，也取得了十分好的效果。

[1] Tran Minh Quan, Junyoung Choi, Haejin Jeong, Won-Ki Jeong. An Intelligent System Approach for Probabilistic Volume Rendering using Hierarchical 3D Convolutional Sparse Coding. IEEE Trans. Vis. Comput. Graph. 24(1): 964-973, 2018.

[2] Joe Kniss, Gordon L. Kindlmann, Charles D. Hansen. Multidimensional Transfer Functions for Interactive Volume Rendering. IEEE Trans. Vis. Comput. Graph. 8(3): 270-285, 2002.

[3] Krishna Prasad Soundararajan, Thomas Schultz. Learning Probabilistic Transfer Functions: A Comparative Study of Classifiers. Comput. Graph. Forum 34(3): 111-120, 2015.

论文报告

交互, 体可视化, 体绘制, 多层次, 机器学习, 科学可视化

[← MovementSlicer:更好的甘特图可视化移动数据中的行为和相遇事件\(MovementSlicer: Better Gantt Charts for Visualizing Behaviors and Meetings in Movement Data\)](#)

[在2D+时间轨迹上评估对时间和速度的感知\(Assessing the Graphical Perception of Time and Speed on 2D+Time Trajectories\) →](#)

评论关闭。

RSS 订阅

功能

登录

文章RSS

评论RSS

WordPress.org

链接

北京大学可视化与可视分析研究小组主页 – PKU Vis Home Page

北京大学可视化研究维基 – PKU Vis WIKI

分类目录

应用

新闻

未分类

活动

研究

论文报告