

Nanocubes for Real-Time Exploration of Spatiotemporal Datasets

在 related work 中吐槽了几个现有的做法：

1. 比如使用分布式集群，但会有延时（我们需要重点考虑这个问题）。
2. 基于 CG 的一些方法，GPU 加速。
3. 采样。
4. 基于每个像素的技术。
5. 比较了令两篇论文中方法：Datavore 和 imMens，认为能比他们处理更大的数据。

Nanocube 的具体算法比较复杂，暂时没细看。

LDAV 调研

浏览 2014 年的 LDAV 的全部文章和 2013 年的部分，未发现有分布式计算与信息可视化结合的文章。

在科学数据上用分布式的太多了，不一篇篇找出来了。

Visual analytics of large-scale climate model data 用了科学数据，但同时使用了信息可视化的方法（平行坐标，散点图），但用的不是分布式

力引导图

算法

http://en.wikipedia.org/wiki/Force-directed_graph_drawing

g

<http://shuli.xianyungu.com/note-guideline-of-the-basic-drawing-methods-1244>

了解了一下目前的做法，需要迭代地找到最优解。

因为每一步的结果会影响下一步，各个节点之间需要知道各自的信息，直接分布式计算延时反而更大。

目前没有想到好的加速方法，还会继续研究。

就目前来看，这并不是一个适合分布计算的问题。

Spark Streaming

了解了一些有关 Spark Streaming 的东西，下周试着写一写。

Spark 1.2 起支持 Python。（改用 Scala 或 Java，或更新 Spark 至 1.3.1）