

Mapwingis:

首先我先介绍一下项目背景:

组内有段时间是研究大图分割与分类的, 研究中分割分类的结果图是用shapefile矢量文件进行保存的。大概的研究方法是把大图切分成小图进行分割, 然后把分割结果逐个追加到shapefile中, 最后我们得到了一个2.5G的矢量图片。此时的问题就是如何打开并显示2G的大图, 组内自己的打开程序只能打开30Mb以下的,

调查市面上主流的GIS的map程序, 其中mapwingis只能部分打开绘制, 无法完全打开, qgis可以完全打开, 但是打开时间要40s并且每次操作还要重复这40s的流程, 于是就决定自己来解决这个问题, 整个项目都由我一人负责。

然后我来介绍一下项目的难点:

第一是: 如何读取2G的矢量数据, 为了达到高效操作的效果, 这2G矢量数据肯定是不能完全加载的。

第二是: 如何加快显示流程中的绘制速度, 每个矢量图形在读取点的信息后都需要请求window系统资源进行绘制。

我的思路是: 2G的矢量文件, 在显示的时候, 肯定不需要全部的信息, 部分信息就足以完成显示, 比如小于一个像素的, 直接在相应位置绘制点, 就不用继续读取其点线集的详细数据。只需要读取当前大于一个像素的矢量图形数据。

就像我们显示地图一样, 在比例尺较大的时候肯定看不到街道, 在比例尺较小的时候也不需要除屏幕内的其他地图部分的信息。但是地图这部分使用的大部分是栅格, 这里使用的是动态金字塔技术, 我们组内之前也对此技术做了解决, 这里与矢量文件的处理方案不同。

我的工作: 首先我详细的分析了shapefile的文件格式。 (展示文件结构) 关键点是每个矢量图形数据都有最小外接矩形的信息, 但是这种方案的问题是还是需要遍历整个shapefile, 虽然此时可以进行绘制但是也需要40s与qgis的绘制时间大致相同, 这肯定是无法满足使用需求的。

在遇到这个问题时, 我分析了一下, 首先我虽然读取的是部分2g信息, 但是是分散在文件中的, 不是连续存储的, 这极大的影响了效率。

其次我对图像的筛选是基于当前区域, 和当前显示的比例尺的。在空间上是具有相关性的。

于是我想到了把读取出来的每个矢量图形对他进行建树, 于是我在这里引入了R树, R树是B树从1维到二维的推广 (r树图)。

这里的总体思路是, 在第一次打开的时候对该矢量文件建立R索引树, 然后保存在本地, 使得该文件下一次打开可以直接使用, 无需再次创建。这个建树时间18s。

这里的R树为什么能加快我们的筛选: 举例 (如图):

可见数据集和不可见数据集

最后一次性的申请系统资源GDI 依次绘制线段即可。

回头解释一下整个流程

这里的R树在我们快速操作所有矢量图形都是有用的。

然后我仔细的解读了mapwingis的文件打开程序，发现了他无法完全绘制的bug，主要是有符号和无符号类型混用导致的。