## 试题6 基于DP算法线状要素图形化简轨迹数据压缩算法

## （原题名称：轨迹数据压缩算法）

手机、汽车、船舶等都会产生大量的轨迹数据，这些数据常有很多冗余，为了便于数据存储与计算，需要对原始数据进行压缩。道格拉斯-普克（Douglas–Peucker ，DP）算法是常用的轨迹压缩算法，对大量冗余的图形数据点进行压缩以提取必要的数据点。如图1所示，通过DP算法对原始数据抽稀，如将原17个点抽稀为4个点。



图1 数据压缩算法示例

### 一、数据文件读取（20分）

编写程序读取“轨迹数据.txt”文件，数据内容和格式如表1所示。

表1 数据内容

|  |  |
| --- | --- |
| 数据内容 | 数据格式 |
| P0,107.605,137.329  P1,122.274,169.126  P2,132.559,179.311  P3,153.324,184.276  P4,171.884,174.654  P5,186.408,168.634  P6,196.566,145.204  P7,200.549,127.877  P8,211.391,118.179  P9,216.318,116.547  P10,225.197,122.796  P11,231.064,135.459  P12,240.835,143.398  P13,254.630,144.933  P14,265.055,158.761  P15,271.004,159.660  P16,274.474,173.979 | 点名，x坐标分量（m），y坐标分量（m） |

### 二、算法实现（60分）

1. 点到直线的垂直距离（20分）。

计算P()到直线 的垂直距离d为：

 （1）

2. DP算法（40分）

该算法实现抽稀的过程是：

（1）将曲线的首末点连成一条直线，求曲线上所有点到直线的垂直距离，并找出最大距离值dmax。

（2）用dmax与事先给定的阈值D相比：若dmax<D，则将这条曲线上的中间点全部舍去，则该直线段作为曲线的近似，该段曲线处理完毕。

（3） 若dmax≥D，保留dmax对应的坐标点，并以该点为界，把曲线分为两部分，对这两部分重复使用该方法，即重复（1），（2）步，直到所有dmax均小于D，即完成对曲线的抽稀。

显然，本算法的抽稀精度与阈值相关，阈值越大，简化程度越大，点减少的越多，反之，化简程度越低，点保留的越多，形状也越趋近于原曲线。

### 三、计算结果报告（20分）

编程输出阈值为5.0和8.0的压缩结果。

### 四、参考答案

--------压缩结果（阈值：5.000） -----

P0,107.605,137.329

P1,122.274,169.126

P3,153.324,184.276

P5,186.408,168.634

P7,200.549,127.877

P9,216.318,116.547

P12,240.835,143.398

P13,254.630,144.933

P15,271.004,159.660

P16,274.474,173.979

--------压缩结果（阈值：8.000） -----

P0,107.605,137.329

P1,122.274,169.126

P3,153.324,184.276

P5,186.408,168.634

P7,200.549,127.877

P9,216.318,116.547

P16,274.474,173.979

程序运行界面如图1所示，显示不同阈值（5.0和8.0）的压缩计算结果。

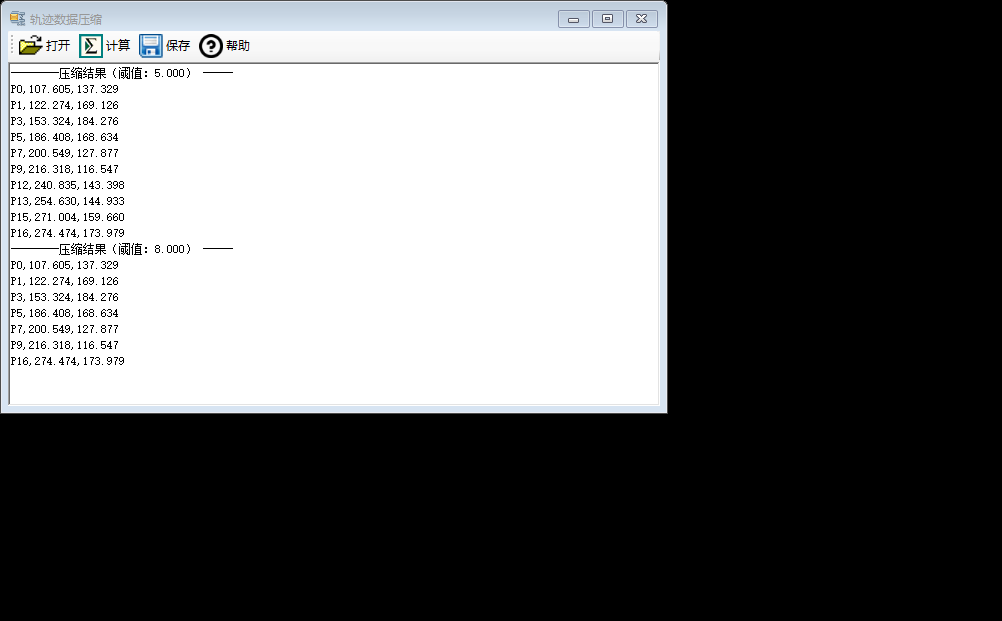


图1 用户界面示例