


轨迹相似性（曲线相似性）

目录

- 曲线相似性分析方法调研
- 基于点的距离
 - 欧式距离 ED
 - 动态时间规整 DTW
 - 最长公共序列 LCSS
 - 编辑距离 EDR
- 基于性状的距离
 - Frechet Distance
 - Hausdorff Distance
- 基于分段的距离
 - One way Distance OWD
 - Locality in-between PolyLines LIP
- 基于特定任务
 - CATS(基于线索感知的轨迹相似度)
 - TRACCLUS(轨迹聚类)
 - NEAT
 - 语义轨迹

 本页面由姚宇辰于2022/04/25迁移自wiki CPT空间

曲线相似性分析方法调研

1	分类	算法	优点	缺点
2	基于距离的	欧氏距离	线性计算时间	轨迹长度要相同且对噪音点敏感。
3		DTW	对轨迹长度无限制，且效果较好；	对噪音比较敏感，离群点也会对结果造成较大影响
4		LCSS	噪音点的偏离没有与其相近的轨迹点故不会被计算在最终结果内，这一步骤有效对抗噪音	算法的最小距离阈值e不好定义
5		EDR	对噪音有一定的鲁棒性，EDR比LCSS更精确	阈值不好定义，对噪音点敏感
6	基于形状的	Hausdorff距离 条件： 两条轨迹之间点的个数不能相差太多，	计算两条轨迹之间最近点距离的最大值，关注形状	对噪音敏感
		Frechet距离	基于动态规划的思想	对噪音敏感

7		discrete Fréchet distance (DFD)		
8	基于分段的方法	单向距离(OWD)	OWD距离的基本思想基于两条轨迹围成的面积，当面积大，说明轨迹之间距离较远，相似度就低；相反，若围成的面积为0，则说明两条轨迹重合，相似度最高。	
		Locality in-between PolyLines LIP多线位置距离	当某区域面积的周长占总长比重大时权重也自然就大；当Area均为0时，说明两条轨迹重合没有缝隙，LIP距离为0；当Area加权和为大时，则说明两条轨迹之间缝隙较大，LIP距离也就大。此外，权重由区域周长占总长比重大决定，也一定程度对抗了噪音点的干扰。	只作用于2维轨迹
9				

EDR和LCSS的比较:

共同点:他们都是基于点的

EDR计算操作代价时需要考虑未匹配的点

不同点：LCSS代价高，EDR代价低

基于点的距离

欧式距离 ED



要求两条轨迹长度相同，一一对应，



 图片上传失败，请重新上传

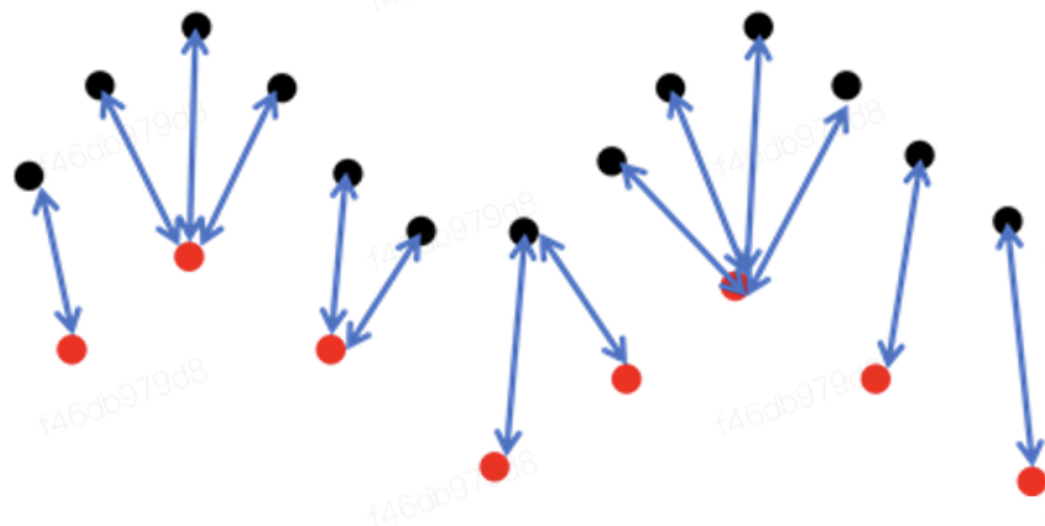
欧式距离的定义简单明了，就是两条轨迹对应点的空间距离的平均值，但是缺点也是很明显的，就是不能度量不同长度的轨迹相似性，而且对噪音点敏感。因此在度量随机事件序列相似度时，缺乏足够精度，其使用范围受到了一定的限制。

欧氏距离用于衡量轨迹相似性的三个缺陷：

1. 不能辨别形状相似性
2. 不能反映趋势动态变化幅度的相似性
3. 基于点距离的计算不能反映不同分析频率的不同。

动态时间规整 DTW

DTW将两条轨迹的点进行多对多的映射，解决数据不对齐问题



动态时间归整算法灵活，对轨迹长度无限制，且效果较好，但是其并未对噪音点进行处理，离群点也会对结果造成较大的影响。

最长公共序列 LCSS

参数 ϵ 是最小距离阈值，两点之间距离小于该值时，被视为同一点。



最长公共子串距离对噪音点进行了处理，即因噪音点的偏离没有与其相近的轨迹点，故不会被计算在最终结果内，这一步骤有效对抗噪音。但与此同时，该算法的最小距离阈值不好定义，还有可能返回并不相似的轨迹。

编辑距离 EDR



在p1处插入一点，将p2替换为p3,和在p5处插入1点，一共3个操作使两条轨迹相等，即对应点距离均小于阈值，故EDR值为3。

轨迹的编辑距离为轨迹相似新度量提供了一种新的思路，其缺陷也很明显，就是对噪音点敏感。同时不好定义阈值。

基于性状的距离

Frechet Distance

，弗雷歇距离就是狗绳距离，即主人走路径A，狗走路径B，各自走完两条路径过程中所需要的最短狗绳长度。





其中， $d(p,q)$ 是两个GPS点的欧式距离， $tr(n-1)=\langle p_1 \rightarrow p_2 \dots p_{n-1} \rangle$ 是轨迹 tr 的长度为 $n-1$ 的子轨迹。

弗雷歇距离为我们提供了一种简单直观的度量相似性的方式，也能达到较好的效果；但可惜的是其并没有对噪音点进行处理，例如若狗的某个轨迹点因为噪声偏离得很远，那么弗雷歇距离也随之增大，这显然是不合理的

要做归一化。

Hausdorff Distance

Hausdorff distance，豪斯多夫距离。简单来说，豪斯多夫距离就是两条轨迹最近点距离的最大值。



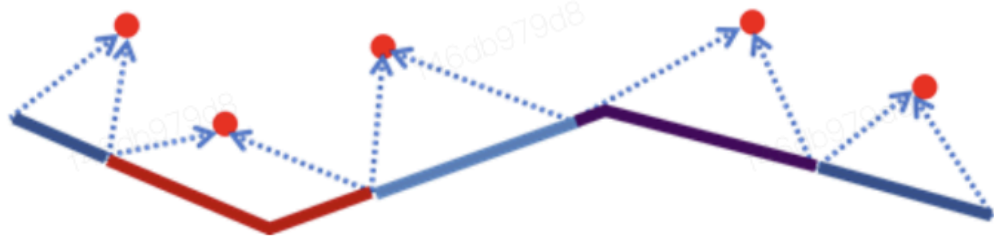


基于分段的距离

One way Distance OWD

OWD距离的基本思想基于两条轨迹围成的面积，当面积大，说明轨迹之间距离较远，相似度就低；相反，若围成的面积为0，则说明两条轨迹重合，相似度最高。

$$d_{owd}(tr_1, tr_2) = \frac{1}{2}(OWD(tr_1, tr_2) + OWD(tr_2, tr_1))$$



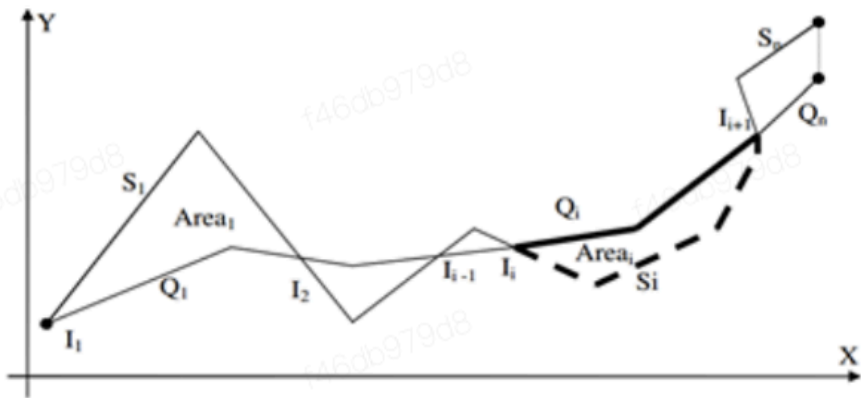
Locality in-between PolyLines LIP

LIP方法很好理解，当某区域面积的周长占总长比重大时权重也自然就大；当Area均为0时，说明两条轨迹重合没有缝隙，LIP距离为0；当Area加权和大的时，则说明两条轨迹之间缝隙较大，LIP距离也就大。此外，权重由区域周长占总长比重大决定，也一定程度对抗了噪音点的干扰。

$$d_{LIP}(tr_1, tr_2) = \sum_{\forall polygon_i} Area_i * w_i$$

其中， li 表示两条轨迹的第*i*个交点，权重 w_i 定义如下：

$$d_{LIP}(tr_1, tr_2) = \sum_{\forall polygon_i} Area_i * w_i$$



基于特定任务

CATS(基于线索感知的轨迹相似度)

由于轨迹在采集的时候可能会存在大量采样点缺失的轨迹段，而对象的同一种运动行为形成的轨迹在空间上和时间上应该都比较接近，因此,对于同一模式的轨迹，将他们的采样点相互补充，得到一条采样完整的轨迹。CATS可以支持局部时间扭曲，对轨迹的采样率和长度都没有要求，并且对噪声具有鲁棒性。

TRACCLUS(轨迹聚类)


通过轨迹聚类找出有代表性的相似轨迹

NEAT

轨迹聚类时,考虑路网因素,分三次聚类

语义轨迹


轨迹序列包含时空语义信息



A survey of trajectory distance measures and performance evaluation.pdf


3.5MB

<https://www.zhihu.com/question/27213170>



曲线相似性分析方法调研(1).docx

548.7KB



Similarity measures for identifying material parameters from hysteresis loops using inverse analysis.pdf

3.6MB