

一种基于阳光的运动轨迹简化算法

茹敬雨¹, 贾子熙², 吴成东²

(1. 东北大学 信息科学与工程学院, 辽宁 沈阳 110819; 2. 东北大学 机器人科学与工程学院, 辽宁 沈阳 110819)

摘 要: 阳光对运动轨迹的影响非常广泛, 高效地估算出轨迹的光照信息, 在简化冗余点的同时保留轨迹的阳光信息至关重要. 本文提出一种基于阳光的运动轨迹简化算法, 结合市内道路的特性解决上述问题. 首先, 提出一种方向模型用以抽象阳光和运动轨迹的关系; 同时, 提出一种与阳光方向相关的运动轨迹简化模型, 并利用遗传算法求取运动轨迹中需要保留的点. 最后, 用美国明尼阿波利斯市的数据进行运动轨迹简化实验, 实验表明运动轨迹的点集数量可以在参数 T_{\max} 的控制下有效地减少.

关 键 词: 智能交通; 轨迹简化; 太阳能; 遗传算法; 时空数据

中图分类号: TP 393

文献标志码: A

文章编号: 1005-3026(2019)08-1070-06

A Sunshine-Based Trajectory Simplification Algorithm

RU Jing-yu¹, JIA Zi-xi², WU Cheng-dong²

(1. School of Information Science & Engineering, Northeastern University, Shenyang 110819, China; 2. School of Robot Science & Engineering, Northeastern University, Shenyang 110819, China. Corresponding author: WU Cheng-dong, E-mail: wuchengdong_neu@outlook.com)

Abstract: Sunshine has played a critical role in trajectory analysis. It is of great significance to design an algorithm that could not only estimate the sunshine information of the trajectory effectively, but also remove the redundant points and keep the necessary sunshine-based information. This paper proposes a sunshine based trajectory simplification algorithm, which considers the urban trajectory character, to solve the challenges of trajectory analyzing in various sunshine related scenarios. On the one hand, an orientation model is established to abstract the relationship between trajectory and sunlight. On the other hand, a sunshine related trajectory simplification model is proposed, and the genetic algorithm is applied to find the points that need to be reserved. Finally, the real-world data of Minneapolis, US, is used for experiment. The results demonstrate that the proposed algorithm can effectively reduce the number of turning points in the trajectories under the control of bound T_{\max} .

Key words: intelligent transportation; trajectory simplification; solar power; genetic algorithm; spatio-temporal data

阳光对人类日常生活的影响不可忽视. 一方面, 阳光每年给地球带来了 3 850 000 EJ 的能量^[1], 其中部分能量被人们在日常生产生活中所利用; 另一方面, 阳光也带来了很多的负面影响, 阳光造成的光污染会增加驾驶员的事故率^[2], 过多的紫外线照射将增加人们得皮肤癌的风险^[3].

由于阳光对人们的日常生活影响巨大, 学者们开始了与之相关的广泛研究. 本篇文章的思路基于这样一个事实——人类相关的运动轨迹非常

容易受光照影响, 同时利用太阳能作为行驶能源或辅助能源的载具及机器人的数量与日俱增, 太阳光对太阳能电池板的照射时长将决定机器人和载具的工作时限, 并影响其运动轨迹规划.

在这些研究当中, 一些学者将重点放在静态环境中的能源收集和设施选址、建设等方面^[4]; 另外一些学者则将重点放在动态移动物体的阳光分析方面^[5], 主要研究了阳光对单一移动物体运动轨迹的影响. 而在实际使用中, 由于系统

模型复杂、点集冗余等问题将降低系统的运行效率,因此如何提高轨迹分析系统的运行效率是本文重要的研究方向.

基于阳光信息的运动轨迹简化对于与阳光相关的轨迹分析系统来说是一个非常重要的环节. 基于阳光的运动轨迹分析既可以帮助人们提高出行体验,又可以提高太阳能电池板的能源收集效率. 而在实际操作中,人们往往会因为运动轨迹点冗余过多,从而造成管理困难、计算量大等实际问题. 基于阳光的运动轨迹简化,可以减少系统的冗余点,简化系统自身需要管理的点的数量,从而提升运行效率.

本研究的挑战在于处理基于阳光分析的运动轨迹时,如何建立一种模型,既可以满足不同的应用,又可以达到需要的精度;根据该模型,可以有效地、可控地简化线段. 为了攻克这一难关,本文提出一种 2D 光照模型,该模型可以有效描述阳光与运动轨迹的光照关系;提出了基于该模型的运动轨迹简化方法,可以有效、可控地简化运动轨迹.

1 研究动机

1.1 运动轨迹分析的挑战

尽管基于阳光的运动轨迹分析有很多好处,但是要实现这样的运动轨迹分析系统面临很多挑战.

模型复杂:不同需求的基于阳光的运动轨迹分析,需要分析不同的阳光方向. 例如,舒适出行需要分析的是阳光来自左侧还是右侧;能源收集需要分析的是阳光的特定角度;安全驾驶需要分析的是阳光来自前方还是后方. 需求不同,则运动轨迹不同,速度不同,阳光相对运动物体的位置也不同,阳光相对不同城市中心点的位置也不同;如果采用完全精确的基于阳光的运动轨迹计算方法对整条路径的光照情况做积分,虽然计算精确,但是计算量将十分巨大.

点集冗余:每一条运动轨迹都是由不同的点组成的,以收集的美国公交车运动轨迹数据为例,全美国共有由 1 360 874 个车站组成的 66 566 条拥有不同形状的交通路线,一共有 31 817 504 个转折点,平均每条运动轨迹就有 478 个转折点. 每有两个相邻的转折点,就会有一个方向的路段,分析如此多的转折点和路段在阳光下的变化情况,是一个非常棘手的问题.

复杂的模型、大量的冗余点、海量的运动轨

迹,导致看似简单的基于阳光的运动轨迹分析,实际上非常复杂.

1.2 市内运动轨迹的特点

由于城市道路的建设特点,在市区内产生的运动轨迹的行驶角度和每次转向的持续时间都有一定的特殊性——都是根据道路的结构,聚集在某些区间内. 市内运动轨迹的行驶角度和每次转向的持续时间的聚集性,是本文设计得以成功的重要保证. 以美国明尼阿波利斯的公共交通轨迹数据集为例,本文首先分析了 2017 年 7 月的周末的 418 990 条运动轨迹,这些运动轨迹包含 2 581 906 个转折点,从中发现市内运动轨迹有碎片化特性.

图 1 是运动轨迹每次转向后的持续时间的累积分布函数 (cumulative distribution function, CDF) 图,可以看到大部分运动轨迹每次转向后的持续时间都小于 10 s. 这是因为数据信息对角度记录得非常详尽,很微小的角度变化就会产生大量新的数据;大部分明尼阿波利斯的街道都非常短,汽车行驶很短的时间就可以通过一条街道;公交车站往往都设置在靠近路口的地方,公交车刚离开车站遇到路口,就会产生新的数据等. 图 2 是运动轨迹的角度分布图,从图中可以看到轨迹的角度多集中在 0° (360°), 90° , 180° , 270° ,这是由于城市的道路结构大多是由矩形组成的.

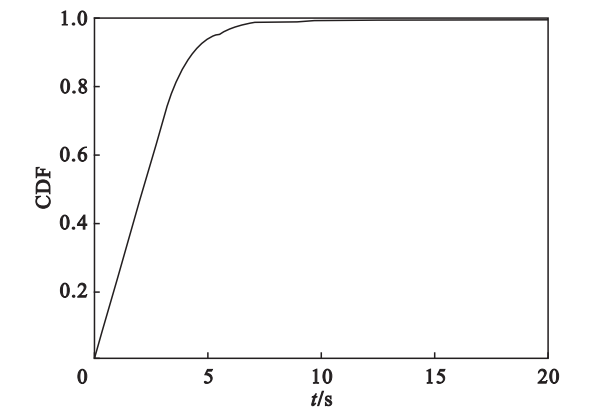


图 1 行驶时间的 CDF 图
Fig. 1 CDF of driving time

1.3 运动轨迹模型的核心思路

本文将利用市内运动轨迹的特性来依次解决 1.1 节提到的两个挑战.

首先讨论复杂模型问题. 就运动轨迹而言,运动轨迹内部的各个路段有不同的方向,根据不同的需求,需要采用不同的研究模型. 本文建立了一个统一的二维光照运动轨迹模型,在满足特定需求的同时,既可以简化运动轨迹内路段,又可以精

确地统计出运动轨迹的光照时间。

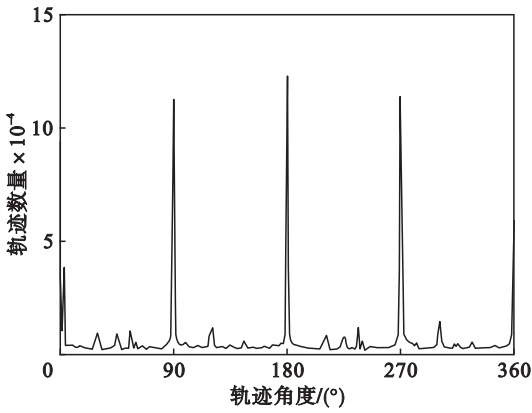


图 2 轨迹角度分布图

Fig. 2 Distribution of directions of trajectories

其次解决点集冗余问题. 通过上面的数据分析可以发现, 一个城市内运动轨迹的角度方向是呈碎片化分布的, 这种碎片化导致每一条运动轨迹都拥有很多点集. 使用运动轨迹内点集数据, 结合相应的光照情况, 可以得到这些运动轨迹的光照信息. 然而, 在这些点集数据中, 不是所有的点都包含了同等重要的光照信息. 本文设计了一种简化法则, 过滤掉包含光照信息少的点, 在宏观逻辑上, 可以减少光照信息分析的计算量.

2 研究内容

2.1 光照模型

在太阳能公交车行驶时, 为了更多地收集能量, 太阳能电池板需要尽量地对准太阳. 但是由于能量所限, 加上汽车的方向不断变化, 实时自适应跟踪太阳的方向并不现实, 所以本文需要周期性地为太阳能电池板设定一个固定的角度, 以达到能源收集期望的最大化. 如何大批量地处理这些运动轨迹, 成为本文的研究内容. 为了更好地展现运动轨迹简化的内容, 本文在传统的光照模型中, 抽象出高效的移动物体的二维方向模型, 如图 3 所示. 假设太阳能电池板的朝向只有左和右, 并依照该假设, 给出如下设计.

假设太阳每天围绕移动物体作顺时针的圆周运动, 如图 3 中虚线圆所示. 定义北为正方向, 同时定义太阳角度 θ_s 为正北到阳光方向的顺时针夹角; 相应地, 定义物体角度 θ_m 为正北到物体前进方向的夹角; θ_o 为正北到运动轨迹反方向的顺时针夹角, $\theta_o = \theta_m + 180^\circ$.

对于任何物体, 用一个三维数据 $(i, \theta_{m(i)}, t_i)$ 和行驶速度 v_i 记录其运动轨迹. 其中 i 代表路段

序号, $\theta_{m(i)}$ 代表物体在路段 i 的前进方向, t_i 代表在路段 i 的行驶时间. 在本文中, 假设运动轨迹匀速行驶, 运动轨迹的长度为 $\sum v_i \cdot t_i$. 所有角度都以正北方向为起点并顺时针旋转, 落入 $[0^\circ, 360^\circ)$ 范围. 以图 4 为例, 该运动轨迹表示为 $(1, 90^\circ, 7s), (2, 180^\circ, 5s), (3, 90^\circ, 6s)$.

当 $\theta_s \in [\theta_m, \theta_o)$ 时, 定义阳光在运动轨迹的右侧, 如图中阴影区域所示; 相反地, 如果阳光满足 $\theta_s \in [\theta_o, \theta_m + 360^\circ)$, 定义阳光在运动轨迹的左侧. 如图 3 所示, 阴影部分代表阳光在运动轨迹右侧, 用 $\xi = -1$ 表示; 白色部分代表阳光在运动轨迹左侧, 用 $\xi = 1$ 表示. 对于任何物体的运动轨迹 $\overline{p_i, p_{i+1}}$, 只要阳光方向确定, 阳光在其左或右的持续时间可以表示为

$$t_{s(i)} = \xi \cdot t_i. \quad (1)$$

以图 4 运动轨迹为例, 当阳光方向为 $\theta_s = 315^\circ$ 时, 光照持续时间为 $(7s, -5s, 6s)$.

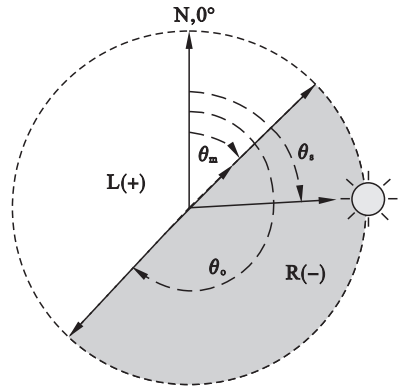


图 3 二维方向模型

Fig. 3 2D orientation model

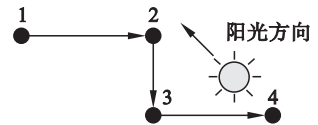


图 4 基于阳光的运动轨迹分析实例

Fig. 4 An example of the sunshine based trajectory

2.2 基于阳光的运动轨迹简化

在阳光运动轨迹分析中, 不是所有的点都承载了相同的阳光信息, 因此通过建立模型, 过滤掉其中重要度低的点.

首先, 用 γ_r, γ_s 分别表示日出和日落的方向. 见图 5a 所示的路段 $\overline{p_i p_j}$, 定义其中 $\theta_{m(i,j)}$ 和 γ_r 之间的角度为

$$\alpha_{L(i,j)} = |\theta_{m(i,j)} - \gamma_r|. \quad (2)$$

相应地, 假设 $\theta_{m(i,j)}$ 和 γ_s 之间的角度为

$$\alpha_{R(i,j)} = |\theta_{m(i,j)} - \gamma_s|. \quad (3)$$

假设太阳每天都作角速度不变的圆周运动,

阳光每天在线段 $\overline{p_i p_j}$ 左侧的概率为

$$P_{L(i,j)} = \alpha_{L(i,j)} / (\gamma_s - \gamma_r); \quad (4)$$

在线段 $\overline{p_i p_j}$ 右侧的概率为

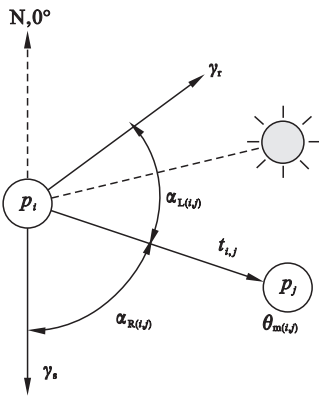
$$P_{R(i,j)} = \alpha_{R(i,j)} / (\gamma_s - \gamma_r). \quad (5)$$

结合线段的行驶时间 $t_{i,j}$,定义该运动轨迹的阳光每日期望时间是

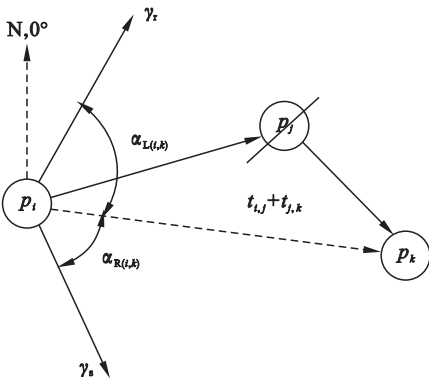
$$E_{t(i,j)} = (P_{L(i,j)} - P_{R(i,j)}) \cdot t_{i,j}. \quad (6)$$

如果一条运动轨迹 $(\overline{p_1 p_2}, \overline{p_2 p_3}, \dots, \overline{p_{n-1} p_n})$ 有包括 n 个点的 $n-1$ 个线段,那么它的阳光每日期望时间是线段阳光期望时间的总和,即

$$E_{t(1,n)} = \sum_{i=1}^{n-1} E_{t(i,i+1)}. \quad (7)$$



(a)



(b)

图 5 基于阳光的运动轨迹简化

Fig. 5 Simplification of sunshine-based trajectory

为了删除运动轨迹中的冗余信息,需要判别数据点的重要性.假设有一条运动轨迹 $(\overline{p_i p_j}, \overline{p_j p_k})$ 由 p_i, p_j, p_k 三点组成,如图 5b 所示.路段初始的阳光期望时间为 $E_t = E_{t(i,j)} + E_{t(j,k)}$.一旦删除点 p_j ,运动轨迹将变成 $\overline{p_i p_k}$,如图中虚线所示.运动轨迹的持续时间不变,仍然是 $t_{i,j} + t_{j,k}$.从 p_i 到 p_k 的左右角度会变成 $\alpha_{L(i,k)}$ 和 $\alpha_{R(i,k)}$.新的光照期望时间将变为 E' ,变化的阳光期望时间为 $\Delta E_t = |E_{t(1,k)} - E'_t|$.用 ΔE_t 当作删除 p_j 的运动轨迹损耗.

因为每条运动轨迹中除了起始点和终点都可以被删除,所以每条拥有 n 个点的运动轨迹都有 2^{n-2} 种重新组合的方式.用 T_{\max} 来约束简化运动轨迹的误差.本文需要选择一种保留点数最少并同时满足 $\Delta E_t \leq T_{\max}$ 的方案,这个问题是一个 NP 难的问题,所以利用遗传算法来求取其近似解^[6-7].在遗传算法中,用 1 和 0 的模式对运动轨迹点的保留与删除进行编码,并用该方式对简化运动轨迹进行编码.用 $|n|$ 表示保留点的个数,种群的适应度函数如下:

$$\text{Fitness}(x) = \text{lb}(\Delta E_t + 1) + \text{lb}|n|. \quad (8)$$

该算法利用了奥卡姆剃刀原理^[5],是一种准确度和连续性的折中方案^[8-10].

对于任意未处理的有 n 个点的运动轨迹 $T(n > 2)$,先生成一个包含原始运动轨迹的随机种群,种群的数量为 $4 \cdot (n - 2)$.找到种群中适应度最小并同时满足 $\Delta E_t \leq T_{\max}$ 的序列,并将其保存为最优序列 T_{best} .然后用遗传算法中标准的突变、交叉、选择,生成新的种群.当进化代数达到 40 或者最优序列连续 8 代没有更新,或者优化过的序列满足 $|n| = 2$ 时,搜索停止,此时最优序列 T_{best} 中保存的,即为运动轨迹 T 基于阳光的近似最优解.此最优解的约束为 $\Delta E_t \leq T_{\max}$.

3 结果与讨论

3.1 实验数据来源

本文实验数据来自于真实世界中的 General Transit Feed Specification (GTFS) 数据. GTFS 数据定义了一种公开的公共交通数据交换格式,现在全球已经有 1 350 个机构共享了他们的 GTFS 数据.图 6 为美国明尼阿波利斯市区的 GTFS 数据示意图,图中的点集为公交车站的位置信息.

3.2 仿真实验

为了验证算法的可行性,本文用明尼苏达 2017 年 7 月的 GTFS 数据进行实验.该数据共包括运动轨迹 66 438 条,共有 497 064 个数据点,平均每条运动轨迹包括 7.5 个点.图 7 为数据的点集数量与参数 T_{\max} 的关系图.从图中可知,随着 T_{\max} 的增长,点集的数量不断减少,当 $T_{\max} = 10 \text{ s/d}$ 时,点集的数量可以减少 1/2.由于起始点和终点不能被删除,所以在 $T_{\max} = \infty$ 时,每条运动轨迹都只保留起始点和终点,即该数据最多可以简化 73% 的数据.



图 6 明尼阿波利斯市的 GTFS 数据
Fig. 6 The GTFS data of Minneapolis

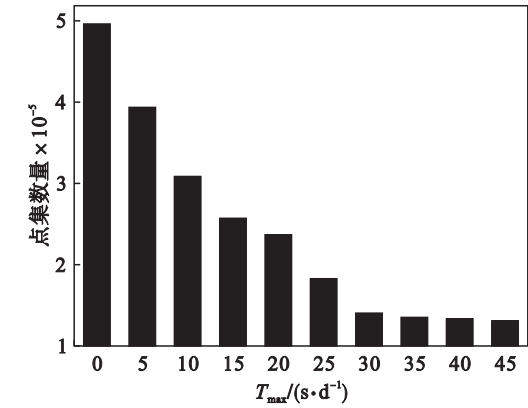


图 7 点集数量与 T_{\max} 的关系
Fig. 7 Relationship of the number of points to T_{\max}

图 8 是 $T_{\max} = 25 \text{ s/d}$ 时,明尼阿波利斯市的 GTFS 运动轨迹数据在相同尺度下简化前后的点密度热力图,此时点集减少约 66%;颜色越深,代表单位区间内点的数量越多.从图中可以看出本文算法非常高效,通过简化,运动轨迹内点的数量大为减少.

4 结 语

本文设计了一种基于阳光信息的运动轨迹简化方法.该方法针对运动轨迹碎片性的特点,利用 2D 光照模型抽象阳光和运动轨迹的关系,设计了一种基于阳光的线段简化方法,解决了模型复杂和点集冗余的问题.真实数据仿真实验结果表明,本文设计的方法可以在可控约束下,以可控的代价有效地减少运动轨迹中点集的数量.基于阳光的运动轨迹分析有广泛的发展前景,本文为其系统的高效运行提供了扎实的理论基础和算法铺垫.

参考文献:

[1] Rhodes C J. Solar energy: principles and possibilities [J]. *Science Progress*, 2010, 93 (1) : 37 – 112.
[2] Abdulaal A, Cintuglu M H, Asfour S, et al. Solving the multivariant EV routing problem incorporating V2G and G2V options [J]. *IEEE Transactions on Transportation Electrification*, 2017, 3 (1) : 238 – 248.
[3] Prochaska J O, Velicer W F, Redding C, et al. Stage-based expert systems to guide a population of primary care patients to quit smoking, eat healthier, prevent skin cancer, and receive regular mammograms [J]. *Preventive Medicine*, 2005, 41 (2) : 406 – 416.

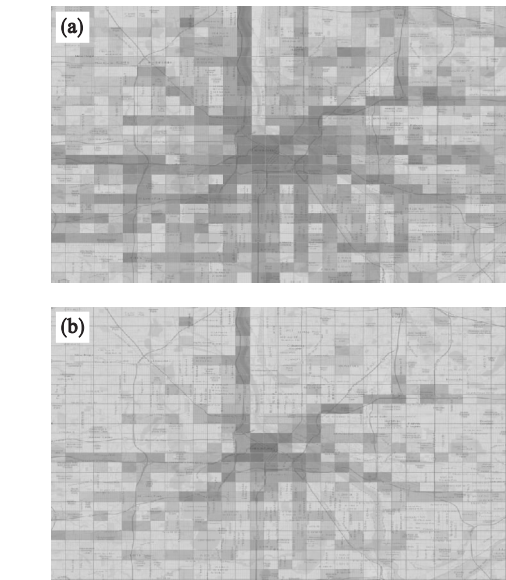


图 8 简化前后点密度热力图对比
Fig. 8 Comparison of heatmaps before and after simplifying points
(a)—简化前热力图; (b)—简化后热力图.