

利用随机图语法的地理视频运动要素解析

丰江帆¹ 宋 虎¹

1 重庆邮电大学计算机科学与技术学院,重庆,400065

摘要:提出了一种利用随机图的地理视频流解析方法,用于对地理视频中的运动要素进行相关性分析。利用运动要素在时空上的相关性和连续性,得到一种反映视频流中运动要素空间关联的时序结构,提供了一个将语义信息与运动信息有效整合的思路。实验证明,该方法不仅能够动态、直观地描述地理视频流中运动要素的空间关系,而且为地理视频场景的语义描述提供了新的思路。

关键词:随机图;地理视频;视频流;解析

中图法分类号:P208

文献标志码:A

视频GIS是当前地理信息科学的研究热点之一,即通过对编码获取的具有空间定位信息的地理视频数据进行传输和管理,并在单帧影像解析基础上进行空间量测和实体三维建模的技术系统^[1]。自从Gimblett在1989年提出将数字视频与空间定位信息相结合以来,国内外从地理视频编码^[2-3]、地理视频组织与管理^[4-5]、视频GIS应用等方面展开了大量研究。

目前的视频GIS均直接采用面向通信或广播的编码方法,没有针对任务本身进行特殊的编码,没有表达数据包含的语义信息,不能适应监控等应用领域对地理视频进行实时解析的需要。虽然地理视频中丰富的空间信息为视频的智能处理提供了解决途径,但由于空间信息固有的不确定性和不完整性,通常只能反映目标的部分知识。如何从错综复杂的地理视频数据中高效地挖掘信息并学习知识,是一个亟待解决的问题。地理视频中的运动要素在视觉中存在位置、形状、大小、速度及其他随时间变化的属性。在运动过程中,运动要素之间存在着相互作用,如人群在运动过程中的自组合状态等。了解这些属性的变化规律,对地理视频的描述与解析具有重要意义。

图是最常用的抽象数据结构,其语法和语义比树、线性表等数据结构更复杂,表现能力更具一般性^[6],适合表现地理视频中要素的空间关系。在连续视频序列图像中,某些运动要素在整体规则运动的过程中伴随着随机运动的出现,运动过

程中的空间关系会随着运动要素的动态性、随机性的变化而不断演化,并且通常情况下运动要素可能只能维持一种空间关系。因此,本文通过学习上下文相关的随机图语法来描述地理视频中运动要素之间在拓扑结构上的相互影响与相互作用关系,并对地理视频流的内容进行了解析。

1 运动要素的语法构造

1.1 实体对象运动要素的描述语法

视频运动要素(video motion element)主要指能够被视觉清晰标识且重要的实体对象,如行人等。目前,针对运动要素的描述主要基于纹理特征、颜色等,难以支持对要素的定义、行为解析以及行为理解。描述运动要素的动态特性,首先需对运动要素的相关概念进行定义。

定义1 状态(State)。状态是对运动要素所有属性的一种抽象,是运动要素在某一时刻所处的状况或活动的一种静态描述。 $State = \{Appear, Move, Stop, Disappear\}$,指地理视频中某一运动要素在空间约束范围内自身呈现的基本状态,包括出现(Appear)、消失(Disappear)、运动(Move)、静止(Stop)等基本状态描述信息。具体描述如下:

1) Appear:在特定地理边界内新出现的、有别于其他已存在的运动要素,本文将其状态称为出现(Appear),Appear实例作为运动要素状态

描述的第一个实例。

2) **Disappear**:与 **Appear** 定义相反,将在地理边界特定区域内消失的或在特定时间范围内无法跟踪的运动要素的状态称为消失(**Disappear**),并取消对该运动要素的标记和跟踪计算,**Disappear** 实例作为运动要素状态描述的最后实例。

3) **Stop**:一个静止(**Stop**) **S** 是由三元组 $S = (\text{Area}(S), \zeta_{\min}(S), \zeta_{\max}(S))$ 构成,其中 $\text{Area}(S)$ 表示 **S** 的平面空间区域, $\zeta_{\min}(S)$ 、 $\zeta_{\max}(S)$ 分别表示 **S** 的最大时间阈值和最小时间阈值。**Stop** 将运动要素在一定平面区域内的运动或停留均视为静止。

4) **Move**:**M**(**Move**)是指空间约束范围内,运动要素的持续运动轨迹中连接其他3种基本状态的元素的统称,**Move** 将其他3种基本状态的实例连接起来,形成一个由 **Appear**、**Stop** 和 **Disappear** 组成的有序序列。

定义2 行为属性(**Behavior Attribute**)。对单一运动要素的行为描述主要包括空间位置(**Location**)和速度(**Speed**)等属性信息。**Location** 可定义为元组 $\text{Location}(\text{Object}) = (X_i, Y_i, T_i)$,表示在时间点 T_i 对象 **Object** 的空间位置关系为 (X_i, Y_i) , X_i 、 Y_i 分别为二维平面坐标中水平坐标值和垂直坐标值。速度定义为 $\text{Speed}(\text{Object}) = \{S_{\text{Value}}, S_{\text{Vector}}, T_i\}$,表示在时间点 T_i 对象 **Object** 的速度数值为 S_{Value} 、速度方向为 S_{Vector} ,其中 S_{Vector} 是平面区域的单位向量。

定义3 关系(**Relation**)。关系是指在相同时间子空间内,描述两个运动要素相互影响的一种关联关系。 $\text{Relation} = (\text{Object}_i, \text{Object}_j, T)$,表示运动要素 Object_i 和 Object_j 在时间子空间 T 建立关系,其中 T 为一维时间坐标。针对已建立 **Relation** 的两个运动要素之间相互作用的描述,可使用概率 P 表示, $P \in [0, 1]$,且 P 的数值会随着时空因素的影响而动态调整。

定义4 空间关系(**Spatial Relationship**)。包括度量关系、方向关系和拓扑关系。空间关系 $\text{SR} = (\text{Measure}, \text{Direction}, \text{Topology})$,其中度量 **Measure** 表示采用某种度量空间的尺度来描述运动要素之间的关系,如距离等;方向 **Direction** 是指在同一平面参考区域中,两个运动要素之间互为源目标和参考目标的相互指向的对等关系。

定义5 视觉特征(**Visual Feature**)。运动要素的视觉特征描述主要包括颜色(**Color**)、纹理(**Texture**)、形状(**Shape**)、大小(**Size**)等。纹理反

映局部像素形成的结构模式和灰度空间分布规律,颜色和形状等底层特征可以定义和描述运动要素本身。

1.2 实体对象运动要素的描述语法

设定行为(**Behavior**)的状态集合为 **BehaviorState**,其典型元素是 τ ,定义为:

$$\tau ::= \text{Appear} \mid \text{Move} \mid \text{Stop} \mid \text{Disappear}$$

其中,**Appear**、**Disappear**、**Move**、**Stop** 分别表示运动要素的4种基本状态描述。

相互作用(**Interaction**)是视频运动要素之间联系的一种表现形式,是指两个行为状态实例之间产生关系(**Relation**)的过程中相互产生的影响或共同作用。建立相互作用关系的条件是两个行为状态实例之间同时存在相互关联关系。可定义为五元组:

$$\text{Interaction} = \{\text{Object}, \text{BehaviorState}, \text{SR}, T, \text{Rule}\}$$

式中,**Interaction** 表示在时间子空间 T 和空间关系 **SR** 影响下,运动要素 Object_i 和 Object_j 之间的相互作用关系描述。**Object** 的行为状态可以是 **BehaviorState** 集合中任一状态实例,**Rule** 包含相互作用产生规则和优化更新规则。因此,相互作用关系的量测包括时间因素和空间因素。

由于在任意时刻 T_{i+1} ,视频运动要素所构成的空间关系总是与前一时刻 T_i 的空间关系紧密相关,因此,这种空间关系演化过程可定义为在时间子空间上的马尔可夫链,即:

$$P_T\{G_{t+1} \mid G_t, G_{t-1}, \dots, G_0\} = P_T\{G_{t+1} \mid G_t\}$$

针对空间关系 **SR** 的度量、方向和拓扑要素,已建立关系 **Relation** 的运动要素之间相互作用的度量值 P 可由平面空间距离(**Distance**)、速度 V 和方向角以及当前时刻 T 的拓扑关系决定。

2 实体对象运动要素的描述语法

随机图 $G = (V, E, \Omega)$ 定义为一个三元组,其中,以 V 为顶点集、 E 为边集的图 G 的边集定义在概率空间 Ω 中,即

$$P(e_{ij} \in E) = P_{ij}, \text{ 且 } P_{ij} \in (0, 1), \sum P_{ij} = 1$$

随机图 G 的各条边均是相互独立的,即任意两个建立关联关系的顶点独立地以概率 p 相连接。由于运动要素在运动过程中的空间关系会随时间因素而动态变化,因此使用随机图对其运动状态和相互作用关系进行描述。上下文相关的随机图语法定义在一个五元组上:

$$G = (S, V_N, R, \delta, Ch)$$

式中, S 为根节点, 表示一个地理视频中语义事件的初始节点, 在一个视频事件发展序列中仅有唯一的 S 节点; 运动要素节点 $V_N = \{V_1, V_2, V_3, \dots\}$ 包含特定空间区域内出现的所有运动要素; R 表示随机图 G 的演化过程与规则; δ 表示随机图的状态转移函数; 随机子图内聚度 Ch (Cohesion) 表示随机子图中运动要素群体内部耦合度。

随机图中的运动要素节点 V_i 定义如下:

$$V_i = (Index, Time, Location, Speed, Interaction, SR, VF, State)$$

表示标记为 $Index$ 的运动要素 V_i 在时间点 $Time$ 所处的运动状况及相互作用关系等信息。其中, $Location$ 和 $Speed$ 分别表示在平面区域内运动要素 V_i 的位置坐标和速度; $Interaction$ 是运动要素之间相互作用关系的描述; $SR = (Measure, Direction, Topology)$ 是运动要素之间存在的空间关系; 视觉特征 VF 表示时间点 $Time$ 运动要素的 $Color$ 、 $Shape$ 、 $Size$ 等底层特征描述信息。状态 $State = \{Appear, Move, Stop, Disappear\}$ 是运动要素 V_i 自身基本状态。

3 实验分析

3.1 相互作用关系

相互作用关系 ($Interaction$) 是运动要素之间彼此存在的相互关联。为准确描述 $Interaction$ 的变化过程, 根据实验视频中运动要素 (监控范围内的人) 的平面空间距离、速度大小和方向等动态信息, 利用动态更新函数实时计算相互作用关系。本实验中, 参数 ρ_1 和 ρ_2 分别设置为 0.15 和 0.85, 置信度 $c(t)$ 初始为 0.36。实验视频片段的 $Interaction$ 计算结果如图 1 所示。

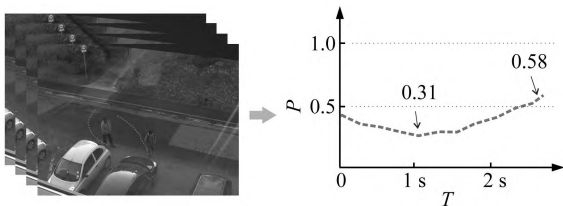


图1 相互作用关系变化过程

Fig.1 Change Process of Interactions

由图 2 可知, 在该视频片段内, $P(t)$ 呈现出先递减后递增的变化趋势。相互作用关系 P 的最小值为第 577 帧时, 取值为 0.31; 最大值为第 615 帧时, 取值为 0.58。初始状态由于两人之间的平面空间距离 $Distance$ 递增, 以及两者的运动方向

的关系, 使得 $Interaction P(t)$ 值不断减小; 随着两者之间平面空间距离 $Distance$ 的减小和相互作用时间的持续增加, 导致 $P(t)$ 值逐渐递增。

3.2 地理视频的描述

针对地理视频中运动要素之间丰富的空间信息, 基于监控感知范围内监控主体的行为状态时序描述, 并根据运动要素在拓扑关系中的相互作用, 利用上下文相关随机图语法, 初步实现了地理监控视频的内容解析。实验视频的第 7 798 帧至 7 845 帧之间的视频片段的随机图演变过程及视频时序发展描述如图 2 所示, 监控范围内出现的人分别标记为 A、B、C 和 D。

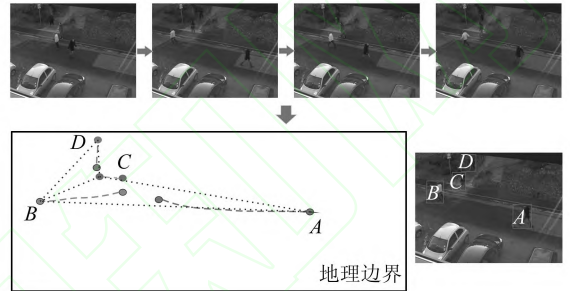


图2 地理视频的时序发展描述图

Fig.2 Description Graph of Timing for Geo-video

由图 3, 随机子图的各运动要素之间是趋于扩散 ($Diffusion$) 的, 随机子图内聚度是递减的。其中, 4 个监控主体 A、B、C 和 D 所形成的相互作用关系更新函数的变化如图 3 所示。

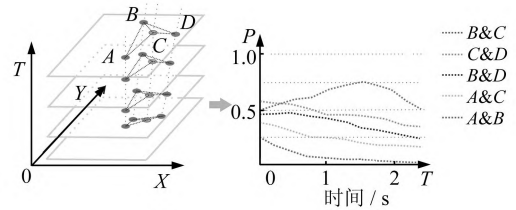


图3 随机子图中相互作用关系

Fig.3 Interactions in Random Subgraph

通过对视频片段的自动分析与计算, 能够构造出视频时序发展描述图, 同时, 也验证了稀疏随机图演化规则算法的正确性和有效性, 有助于实现视频事件理解与异常检测。

4 结 语

实验表明, 利用随机图的地理视频流描述方法是可行并有效的。通过地理视频的智能解析, 构造运动要素的随机图描述, 可以清晰地描绘场景的发展线索; 同时, 可作为地理视频一种有效的非线性索引方式。因此, 在视频监控领域具有重

要的应用价值。

总之,该方法适合从具有相近空间属性的运动要素中挖掘空间语义,形成对地理视频的智能描述。但随机图节点只能模拟具有已知属性的运动要素,当引入具有未知属性的运动要素时,需要建立新的模型以分析新属性与已知属性之间的相关性。此外,进一步的研究还包括复杂属性的描述,例如空间语义、粒度、角色等;局限于度量运动要素的动态关联等。

参 考 文 献

- [1] Feng Jiangfan, Zhu Guanyu. Research of Vehicle Navigation Based Video-GIS[J]. Journal of Korea Spatial Information System, 2009, 11(2):39-44
- [2] Chen Liang, Zhu Changqing, Ren Na, et al. Watermarking Algorithm Using Wavelet Transform for GIS Video Data [J]. Geomatics and Information Science of Wuhan University, 2012, 37(10): 1 256-1 259 (陈亮,朱长青,任娜,等.利用小波变换的视频GIS数据数字水印算法[J].武汉大学学报·信息科学版,2012,37(10):1 256-1 259)
- [3] H wang T H, Choi K H, Joo I H, et al. MPEG-7

Metadata for Video-based GIS Applications [C]. I GARSS, Toulouse, France, 2003

- [4] Kong Yunfeng. Design of GeoVideo Data Model and Implementation of Web-Based Video GIS [J]. Geomatics and Information Science of Wuhan University, 2010, 35(2): 133-137 (孔云峰.地理视频数据模型设计及网络视频GIS实现[J].武汉大学学报·信息科学版,2010,35(2):133-137)
- [5] Paul L. BSc. Linking Spatial Video and GIS[D]. Ireland: Department of Computer Science & National Centre for Geocomputation National University of Ireland Maynooth, 2009:59-70
- [6] Yu Ge, Gu Yu, Bao Yubin, et al. Large Scale Graph Data Processing on Cloud Computing Environments [J]. Chinese Journal of Computers, 2011, 34(10):1 753-1 767 (于戈,谷峪,鲍玉斌,等.云计算环境下的大规模图数据处理技术[J].计算机学报,2011,34(10):1 753-1 767)
- [7] Chen J J, An G C, Zhang S F, et al. A Mean Shift Algorithm Based on Modified Parzen Window for Small Target Tracking [C]. IEEE International Conference on Acoustics, Speech and Signal Processing, Dallas, 2010

Analytical Method for Mobile Elements in Geo-Video Using Random Graph Grammar

FENG Jiangfan¹ SONG Hu¹

¹ College of Computer Science and Technology, Chongqing University of Post and Telecommunications, Chongqing 40065, China

Abstract: Decomposition of a Geo-video stream presents the expression of video from spatial feature set. Although it has been studied widely, spatial relations underlying a scenario are not well understood. Here we use a method which takes advantage of random graph theory to investigate the semantic knowledge in Geo-video, leading to correlation analysis of the target motion elements in a Geo-video stream. We used the connections of target motion elements, both correlation and continuity, to lead to a structure in time series that reveals dues to the event development of the video stream. Furthermore, it provides a method for the effective integration of semantic and campaign information. Ultimately, experimental results show that the method offers a better description of video elements than can be achieved with existing schemes. In addition, the proposed method offers a new way of thinking about the semantic description of the geographic video scenarios.

Key words: random graph; geo-video; video stream; analysis

First author: FENG Jiangfan, PhD, assistant professor, specializes in the video GIS and mobile GIS. E-mail: fengjf@cqupt.edu.cn

Foundation support: The National Natural Science Foundation of China, Nos. 41101432, 41201378; the Natural Science Foundation Project of Chongqing, No. CSTC 2010BB2416.