

引文格式: XIE Xiao, ZHU Qing, ZHANG Yeting, et al. Hierarchical Semantic Model of Geovideo[J]. Acta Geodaetica et Cartographica Sinica, 2015, 44(5): 555-562. (谢潇, 朱庆, 张叶廷, 等. 多层次地理视频语义模型[J]. 测绘学报, 2015, 44(5): 555-562.) DOI: 10.11947/j. AGCS.2015.20140176

多层次地理视频语义模型

谢潇^{1,2}, 朱庆^{3,1}, 张叶廷¹, 周艳⁴, 许伟平¹, 吴晨¹

1. 武汉大学测绘遥感信息工程国家重点实验室, 湖北 武汉 430079; 2. 慕尼黑工业大学地图制图系, 德国 慕尼黑 80333; 3. 西南交通大学地球科学与环境工程学院, 四川 成都 610031; 4. 电子科技大学资源与环境学院, 四川 成都 611731

Hierarchical Semantic Model of Geovideo

XIE Xiao^{1,2}, ZHU Qing^{3,1}, ZHANG Yeting¹, ZHOU Yan⁴, XU Weiping¹, WU Chen¹

1. State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China; 2. Department of Cartography, Technical University of Munich, Munich 80333, Germany; 3. Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu 610031, China; 4. School of Resources and Environment, University of Electronic Science and Technology of China, Chengdu 611731, China.

Abstract: The public security incidents were getting increasingly challenging with regard to their new features, including multi-scale mobility, multistage dynamic evolution, as well as spatiotemporal concurrency and uncertainty in the complex urban environment. However, the existing video models, which were used/ designed for independent archive or local analysis of surveillance video, have seriously inhibited emergency response to the urgent requirements. Aiming at the explicit representation of change mechanism in video, the paper proposed a novel hierarchical geovideo semantic model using UML. This model was characterized by the hierarchical representation of both data structure and semantics based on the change-oriented three domains (feature domain, process domain and event domain) instead of overall semantic description of video streaming; combining both geographical semantics and video content semantics, in support of global semantic association between multiple geovideo data. The public security incidents by video surveillance are inspected as an example to illustrate the validity of this model.

Key words: video GIS; geovideo; semantic model; event; spatiotemporal association

Foundation support: Open Foundation Funded by Key Laboratory of Oasis Ecology of Xinjiang University, China (No.XJDX0201-2013-01); The National Key Technology Research and Development Program of the Ministry of Science and Technology of China (No.2012BAH35B02); The National Natural Science Foundation of China (Nos.41261086;41471320; 41471332)

摘要: 针对公共安全应急响应中支持推理、挖掘和关联分析的地理视频建模难题, 提出了一种显式表达视频变化的多层次地理视频语义模型, 并用 UML 图进行描述。该模型的特点是: 改变了传统视频流整体语义描述方法, 通过面向变化的三域(特征域-行为过程域-事件域)定义地理视频语义的层次结构和数据的层次表达; 在各层次语义描述中将地理环境语义与视频内容语义有机结合, 支持多地理视频数据的关联表示。以公共安全事件监控视频为例阐明了模型的实用性和有效性。

关键词: 视频 GIS; 地理视频; 语义模型; 事件; 时空关联

中图分类号: P208

文献标识码: A

文章编号: 1001-1595(2015)05-0555-08

基金项目: 绿洲生态教育部重点实验室开放课题(XJDX0201-2013-01); 国家科技支撑计划(2012BAH35B02); 国家自然科学基金(41261086;41471320; 41471332)

1 引言

地理视频(geovideo)是包含地理时空信息的视频数据。它具有对地理空间动态、实时和真实感表达的优势,符合人类直观感知和认知特点,是智慧城市与城市安全领域广泛采用的关键信息内容。复杂城市环境中公共安全事件呈现出多尺度流动、多阶段演化以及时空并发与不确定等新特性,公共安全应急响应亟须突破传统分散独立存档和局部解析分析的视频表示与处理方式,实现支持推理、挖掘和关联分析的地理视频建模。

现有地理视频建模研究主要集中在地理信息系统和计算机两个领域。

GIS领域的相关研究成果主要面向视频数据与空间数据的关联表示,其中,空间数据主要采用全球定位系统获取的摄像机空间参考信息。现有研究成果根据关联方式的不同可以分为时间关联模式和位置关联模式两类:①时间关联模式,其特点是以时间戳为基准,通过时间索引实现视频数据与GPS定位信息的同步关联;②位置关联模式,其特点是以相机定位信息为基础进行视频关联。根据实现方法的不同,位置关联模式的研究成果又可分为两类:一类是通过视频与GPS定位信息的实时调制实现二者的关联,代表性成果包括在音频信道上调制GPS信号^[1]以及利用高级流媒体格式(advanced systems format, ASF)实现定位信息、视频音频信息的调制编码^[2];另一类是通过在摄像机定位信息基础上扩展几何、专题、属性、语义标注等元数据实现视频数据与空间数据的关联,代表性成果包括扩展摄像机的地址属性与GPS定位信息^[3-4],扩展相机GPS定位、姿态和成像参数信息^[5-6],扩展成像的视锥体、视景体模型等^[7]。

计算机领域的相关研究,则主要侧重视频内容中的地理对象及其局部时空关系建模。这些对象包括了图形对象和语义对象。根据建模内容的侧重点的不同可分为3类:①侧重离散视频帧记录的实体对象及对象间空间关系表示^[6,8-10],这类研究成果可支持视频图像中对象空间位置的查询以及包含指定地理对象的视频数据查询;②侧重连续视频序列内物理对象及其运动轨迹的时空连续性与时空关系表示^[11-15],这类研究成果可支持连续视频序列中地理对象运动轨迹的空间查询;③侧重连续视频序列内实体(包括具体的物理对象和抽象的角色、事件等)及其属性与时序关系表

示^[16-20],这类研究成果可支持连续视频序列内容的语义查询。在实现方式上,这些研究均采用基于动态图像专家组(moving pictures experts group, MPEG)的可扩展标记语言(extensible markup language, XML)文件来描述视频图像或视频镜头。

综上所述,现有GIS领域的地理视频建模研究主要面向摄像机的空间参考信息,缺乏对视频内容中地理场景的建模和分析。同时,由于传统GIS及其专题扩展模型仍以地理实体和地理过程的几何与属性特征表示为主^[21-23],现有模型也难以有效表示与传统地理空间数据存在明显差异的非结构化地理视频以及视频内容中包含的具有高维语义关联性的地理场景。因此,现有视频GIS系统仍主要将视频数据作为空间数据的属性,独立存储和可视化;系统能力也局限在视频流或视频片段与空间位置的交互检索,难以实现视频数据中监控对象、监控区域与地理环境信息的关联分析。计算机领域的现有研究均着重探讨单一视频帧的图形对象建模和连续视频序列中的小尺度时空关系表示,缺乏对大尺度多视频数据内容间语义和语境的研究。因此,对于非连续和跨区域的监控视频,仅能依靠图像的相似性实现关联性判断,难以支持地理视频数据之间、视频内容与地理环境之间复杂时空关联性统一描述。

地理视频可视为对动态地理环境在特定尺度时空窗口的映射,地理视频反映了地理环境的变化片段。由于这些变化片段对应的地理问题存在复杂的领域综合、动态演化、区域关联和尺度依赖,导致地理视频内容在时间、空间和专题等基本变化要素层次的关联错综复杂^[24];多传感器类型、多视角、多时间粒度与多时空分辨率特性使地理视频和动态地理环境间的映射关系更加多样化,加剧了地理视频内容基本变化要素间关联的复杂度。为此,本文面向公共安全应用,提出一种表达视频内容变化的多层次地理视频语义模型。

2 地理视频语义模型的层次结构

面向地理视频内容变化的显示表示,将地理视频场景变化中的载体、驱动力和呈现模式3个关键因素及其相互关系具体化为具有关联性的地理实体和场景、对象行为和多层次事件对象,并依次抽象为相互关联的特征域、行为过程域和事件域3个层次。为支持复杂地理环境中不同地理视频的关联表示与推理,在各数据层次内容语义的

基础上引入统一时空框架下的地理语义,其概念模型如图 1 所示。

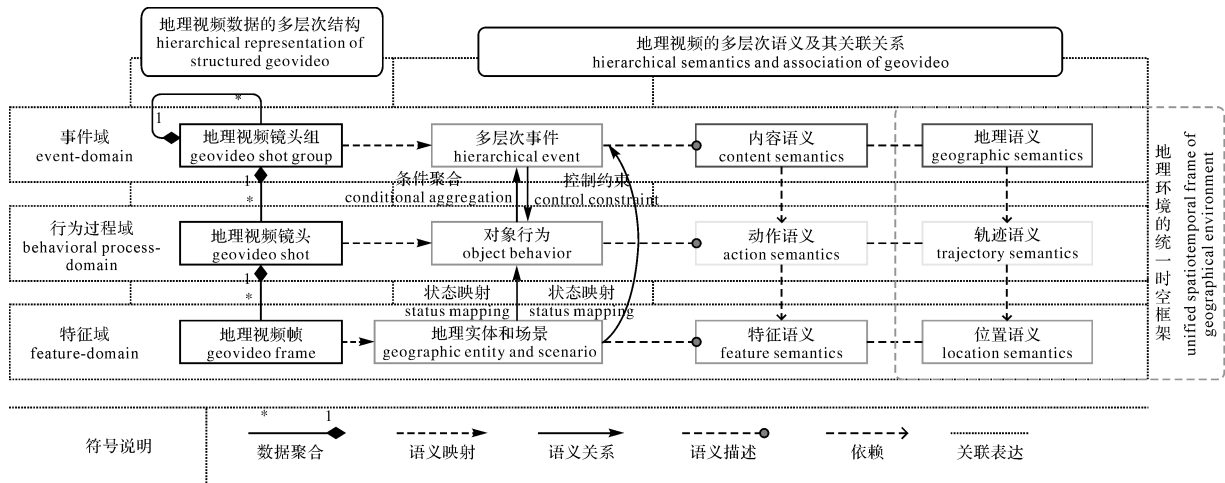


图 1 面向特征-行为过程-事件三域的多层次地理视频语义概念模型

Fig.1 Three-domain-oriented hierarchical semantic conceptual model of geovideo

特征域 (feature-domain)。地理实体和场景 (geographic entity and scenario, O_{ge}) 作为特征对象是地理环境变化的载体,地理实体通常表示具有改变自身状态的行为能力的对象,而场景通常表示为状态相对不变的对象,由条件 (condition)、实例 (instance)、语义 (semantics) 和关系 (relationship) 4 元组描述,形式化为: $O_{ge} = (\{C\}, \{I\}, \{S_{ge}\}, \{R\})$ 。其中, $\{C\}$ 约定了理解和表达 O_{ge} 的时空尺度、结构类型、状态的相对变化性质等条件,如二维图像、三维模型等实例化形式; $\{I\}$ 为 $\{C\}$ 限定下的对象结构,如栅格图像、真三维几何模型等; $\{S_{ge}\}$ 为 O_{ge} 的语义描述,表示为 O_{ge} 的特征参量与附加属性,体现 O_{ge} 具有的变化条件和变化能力,包括对象的时空属性等特征语义 (feature semantics) 及其在地理环境统一时空框架下的位置语义 (location semantics)。其中,位置语义包括:绝对或相对位置描述,位置的拓扑、方位和度量关系,位置的颜色、范围、组成、用途等内涵属性及由此产生的外延规则,位置语义支持对象时空分布的描述与位置关联的表达,其外延规则同时为事件的判定提供基础; $\{R\}$ 表示 O_{ge} 与对象行为和多层次事件的状态映射关系 (state mapping)。

行为过程域 (behavioral process-domain)。对象行为 (object behavior, O_{ob}) 作为过程域对象是地理视频内容变化的驱动力,对应了地理实体的状态、时空关系及属性的变化过程,是地理视频解析与分析的基本单元,由变化流程 (activity process)、关键状态 (key state)、语义 (semantics)

和关系 (relationship) 4 元组描述,形式化为: $O_{ob} = (AP(O_{ob}), KS(O_{ob}), \{S_{ob}\}, \{R\})$ 。其中, $AP(O_{ob})$ 包括对象行为发生的环境描述、表达式或有序离散点描述的对象运动轨迹、对象内部或对象间的关系变化轨迹,记录为表示连续变化的函数/解析式等非线性模型或表示离散变化的线性时间戳模型; $KS(O_{ob})$ 为行为生命周期中一系列重要状态的显示表达,如初始、突变、终止状态等; $\{S_{ob}\}$ 为 O_{ob} 的行为趋势、结果及影响的语义描述,包括描述行为类型与特点等的动作语义 (action semantics) 和建立在地理环境统一时空框架下的轨迹语义 (trajectory semantics)。其中,动作语义强调行为特征的描述 (如:无参照的静止、步行、跑等动作描述,有参照的靠近、通过等动作描述); 轨迹语义侧重对行为的结果及影响的描述,轨迹语义包括轨迹特征语义 (如:走-停-走)、轨迹地理语义 (如:从位置 A 到位置 B、经过位置 C) 以及轨迹关系 (如:相遇、平行等)。同时,对象行为依赖于地理实体和场景,一个对象的特征语义决定了它的行为能力,而位置语义的外延规则约束了行为的可行性,为异常判断和事件推理提供了支持。因此在实际应用中需要针对的典型对象类型,对专题知识中涉及的行为进行预定义和分类; $\{R\}$ 表示 O_{ob} 与多层次事件的条件聚合关系 (conditional aggregation),聚合作用体现在:对象行为需根据其满足的语义关联要素进行语义关联关系的推理构建行为链,进而实现多层次事件的表达。聚合规则为:设 R_U 为专题领域的事件规则库, $\{Ru\}_x$ 为某事

件规则集合, $\{Ru\}_x \subseteq R_U$, 当有 $P(O_{ob})_1, P(O_{ob})_2, \dots, P(O_{ob})_n$, 且 $\bigcup_{m=1}^n \{P(O_{ob})\}_m \mid = \{Ru\}_x$, 那么 $\exists O_{he-x}$, 使 $\bigcup_{m=1}^n \{P(O_{ob})\}_m \in O_{he-x}$ 。

事件域(event-domain)。多层次事件(hierarchical event, O_{he})作为事件域对象是地理视频内容变化呈现模式的抽象描述, 由有序的对象行为链组成, 事件的层次性体现了行为变化的复杂性, 表现为支持不同尺度事件对象因影响、反馈和关联而相互影响而递归聚合为局部小尺度事件、区域中尺度事件和全局大尺度事件, 聚合规则为: 仍设 $\{Ru\}_x \subseteq R_U$, 当 $\{P(O_{ob})\}_1 \in O_{he-1}, \{P(O_{ob})\}_2 \in O_{he-2}, \dots, \{P(O_{ob})\}_n \in O_{he-n}$, 且 $\bigcup_{m=1}^n \{P(O_{ob})\}_m \mid = \{Ru\}_x$, $\exists O_{Er}$, 使 $\bigcup_{m=1}^n \{P(O_{ob})\}_m \in O_{Er}$, 其中 O_{Er} 为事件 O_{he} 的子事件。 O_{he} 由事件规则(rule)、过程集合(process)、语义(semantics)和关系(relationship)4元组描述, 形式化为: $O_{he} = (\{Ru\}, \{P(O_{ob})\}, \{S_{he}\}, \{R\})$ 。其中, $\{Ru\}$ 表示事件判断和推理的一系列规则, 表示为事件模板, 是应用领域事件规则库的元素或子集, 是对 O_{he} 包含有序关联性对象行为更高层次含义的理解和表达; $\{P(O_{ob})\}$ 表示构成事件的一系列有序过程集合; $\{S_{he}\}$ 是对事件知识的语义描述, 包括事件的内容语义(content semantics)以及事件发生环境的地理语义(geographic semantics); $\{R\}$ 为事件对象对过程对象的控制约束关系(control constraint), 控制约束作用体现在: 根据多层次事件演进阶段、因果规律、发展趋势和应用层次等内在变化阶段抽象出的反馈、原因、影响、包含等语义关系, 形成对视频内容中对象行为语义关系判断, 如前后两者反馈不同演进阶段的顺序行为; 前后两者呈现因果性触发条件的条件行为; 前者对后者施加控制约束的约束行为; 前后两者属于不同层次, 且前者高于后者的聚类行为。

3 地理视频数据的多层次表达

以视频内容的变化即人的行为和周围环境变化为基本单元, 将地理视频分为地理视频帧、地理视频镜头和地理视频镜头组分为3个粒度层次, 各层次的形式定义和结构特点如下。

地理视频帧(geovideo frame, O_{gf}), 地理视频数据的最小结构粒度和数据变化的解析单元, 形式化为: $O_{gf} = (\{C\}, \{I\}, \{S_{gf}\}, \{R\})$ 。其中, $\{C\}$

表示编码格式、码率、帧率、分辨率等; $\{I\}$ 为与编码格式对应的静态图像对象, 其中, 状态相对变化的地理实体和状态相对静态的地理场景分别实例化为分离提取前景图像和背景图像; $\{S_{gf}\}$ 包括图像的摄像机方位、姿态、时刻、成像参数等图像物理特征描述, 以及可选的领域相关的图像分割规则等外部语义描述, 图像内容蕴含的语义对象等内部语义描述, 对应了特征域对象; $\{R\}$ 表示 O_{gf} 与地理视频镜头、地理视频镜头组的状态映射关系。

地理视频镜头(geovideo shot, O_{gs}), 结构化地理视频数据的变化单元。面向过程域的 O_{gs} 表示中, $O_{gs} = (AP(O_{gf}), KS(O_{gf}), \{S_{gs}\}, \{R\})$; 其中, $AP(O_{gf})$ 为状态映射关系下, 基于数据相似性划分的连续地理视频帧序列, 该序列中的帧有相同的实例化条件 $\{C\}$ 和相似语义项 $\{S_{gf}\}$ 取值; $KS(O_{gf})$ 为 $AP(O_{gf})$ 中具有语义项峰值的一个或具有相互间最大不相关性的多个地理视频帧; $\{S_{gs}\}$ 包括镜头对应的摄像机外部语义描述以及镜头内容中所能解析出的一系列对象行为, 对应了行为过程域对象, $\{S_{gs}\}$ 是 O_{gf} 在时间维主导下所能表达的更高维度的语义信息, 如从某时刻到生命周期的语义表达等; $\{R\}$ 表示 O_{gs} 与地理视频镜头组的条件聚合关系, 由 $AP(O_{gf})$ 决定。地理视频镜头 O_{gs} 是地理视频语义建模中地理实体行为解析与表示的基础。

地理视频镜头组(geovideo shot group, O_{gsg}), 呈现专题规则下变化的形成与发展的有序地理视频镜头集合。相对于传统由单摄像头记录的物理上连续的视频对象, 地理视频镜头组可视为一个虚拟视频对象, 支持来自不同摄像机但变化逻辑连续的地理视频镜头。面向事件域的 O_{gsg} 表示中, $O_{gsg} = (\{Ru\}, \{P(O_{gs})\}, \{S_{gsg}\}, \{R\})$; 其中, $\{Ru\}$ 为支持地理视频镜头关联的规则库子集; $\{P(O_{gs})\}$ 为满足 $\{Ru\}$ 条件聚合的地理视频镜头集合, 满足 $O_{gs} - O_{gsg}$ 以及 $O_{gsg} - O_{gsg}$ 递归嵌套的聚合规则, 不同于 O_{gf} 与 O_{gs} 间基于数据相似性的划分, O_{gs} 与 O_{gsg} 的聚合将基于地理视频的语义关联推理实现; $\{S_{gsg}\}$ 为对 O_{gsg} 内容语义的综合描述, 是 $\{P(O_{gs})\}$ 中 $\{S_{gs}\}$ 关联表达后所能反映的更高层次主题含义, 对应了事件域对象; $\{R\}$ 表示 O_{gsg} 对 O_{gs} 的控制约束关系, 这些关系可分为如下两大类: ①地理视频场景间的时空关联, 包括多视点关联, 多视角关联和多分辨率关联等; ②地理视频场景对象间的时空关联, 包括典型的 GIS 对象间的时空拓扑关系以及事件间反馈、原因影

响与包含的语义关系。实际应用中,不同地理视频镜头、地理视频镜头组的语义关联关系,需通过对视频内容的推理和判断得到;其中,地理视频数据不同层次的语义描述是推理和判断的依据。此外,镜头组的实例依赖于多层次事件的实例,对多层次事件的表达是地理视频镜头与地理视频镜头组 O_{gs} 关联的主要目的。记录了某特定事件的一组镜头对象才聚合表示为对应于该事件的镜头组,镜头组的聚合除了反映不同时空尺度的事件,还可表达那些内容存在空间维度的重叠,且因摄像机成像的不同时空分辨率而对同一地理事件不同细节层次表达的多镜头。

基于统一建模语言(unified modeling language, UML),设计了多层次地理视频语义模型类图结构。其中,地理视频数据的 3 个粒度:地理视频帧、地理视频镜头和地理视频镜头组分别表达为派生于父类 GeovideoStructure 的 3 个核心子类 CGeovideoFrame、CGeovideoShot 和 CGeovideoShotGroup。为了支持这些对象类在实际数据和成像信息上的描述,设计了表达摄像机元数据和实际图像编码的 CCamera、CViewpoint、CSimpleImage 等 10 个附加类,这些信息在表达实际地理视频数据的基础上,还是判断地理视频场景间的时空关联的基础。更重要的是,为了支持地理视频的语义描述,将与地理视频数据相映射的 3 个地理视频语义层次:地理实体和场景、对象行为和多层次事件分别表达为派生于父类 CInnerSemanticObject 的 3 个核心子类 CGeoEntity、CObjectBehavior、CHierarchicalEvent;同时,设计了支持这些语义对象内容描述的特征语义类(CPropertySemantic)/动作语义

类(CActionSemantics)以及支持地理语义描述的位置语义类(CLocationSemantics)/轨迹语义类(CTrajectorySemantics)。引入以 CGeoReferenceTrans 为父类的统一时空语义框架作为地理语义类的重要成员变量,它是支持多地理视频内容和地理环境语义映射基础。

4 实例分析

本文以博物馆公共安全监控的多路视频数据为例,对其进行分析,利用文中提出的地理视频语义模型实现监控视频数据的元数据描述和关联表达。实例数据选用包含 20 路视频的监控网,包括编号依次为 C-01 至 C-20 的相机视域及分布情况,分别对应了针对展品、展厅和楼层等不同尺度的监控范围。

博物馆公共安全监控视频数据的三域要素如表 1 所示:①在特征域,以地理视频帧为分析单元,从变化的地理视频帧中解析出“人”(员工/来访者)、“展品”和“地理环境部件”(室内空间/建筑结构)3 大类特征要素,地理环境部件是基于城市地理标识语言 OGC CityGML 和室内多维位置信息标识语言 OGC IndoorGML 的国际标准,包含室内地理位置的空间关系及其包含、联通等语义关系,用于支持位置语义和轨迹语义的表达;②在行为过程域,根据图像像素域的视觉、光学等物理特征以及图像压缩域编码规则的相似性阈值划分出包含人员行为变化的地理视频镜头,并解析其中人员的主动性行为(分别从动作模型和轨迹地理语义两方面描述)和展品的被动性行为;③在事件域,面向文物(展品)突发安全事件,将与事件相关的镜头聚合为有序地理视频镜头组。

表 1 博物馆公共安全监控视频的三域要素

Tab.1 The three-domain of geovideo in museum security monitoring

	地理视频数据的结构	地理视频的语义
特征域	$\{O_{gf} \text{当 } \Delta(O_{gf}, \{I\})! = \Phi\}$	$\{O_{ge} O_{ge} \in \text{人} \vee \text{展品} \vee \text{地理环境部件}\}$ +人 = $\{p p \in \text{员工} \vee p \in \text{来访者}\}$ +展品 = $\{e e = \text{ID}_1, \text{ID}_2, \dots, \text{ID}_n, n \in N\}$ +地理环境部件 = $\{c c = \text{建筑结构} \vee \text{室内空间}\}$
行为过程域	$\{O_{gs} \text{当 } \Delta(O_{gs}, AP(O_{gf}), \{\{C\}, \{S_{gf}\}\}) < \text{相似性阈值}\}$	$\{O_{ob} O_{ob} \in \text{主动行为}\langle p \rangle \vee \text{被动行为}\langle e \rangle\}$ +主动行为 $\langle p \rangle = \{(\text{行为模型})\langle p \rangle \wedge \langle p \rangle \text{轨迹语义}\langle c \rangle\}$ +行为模型 $\langle p \rangle = \{\text{跑}\langle p \rangle \vee \text{步行}\langle p \rangle \vee \dots, \text{等}\}$ +轨迹语义 $\langle c \rangle = \{\text{经过}\langle c \rangle \vee \text{从}\langle c \rangle \text{到}\langle c \rangle \vee \text{徘徊于}\langle c \rangle \vee \text{停留于}\langle c \rangle \vee \text{向}\langle c \rangle \text{靠近} \vee \text{从}\langle c \rangle \text{离开} \vee \text{来到}\langle c \rangle \vee \text{进入}\langle c \rangle \vee \text{从}\langle c \rangle \text{出} \dots, \text{等}\}$ +被动行为 $\langle e \rangle = \{(\text{被移动})\langle e \rangle, \langle p \rangle\}$
事件域	$\{O_{gse} \text{当 } O_{gse}, \{P(O_{gs})\} = \text{聚合规则} \text{事件模板}\}$	$\{O_{he} O_{ob} = \text{文物安全突发事件(SEHR)}(\langle p \rangle, \langle e \rangle)\}$ +SEHR = $\{\text{实施}\langle \{p\}, \langle e \rangle \rangle \wedge (\text{infiltrate}\langle \{p\} \rangle \text{追击}\langle \{p\} \rangle \text{逃跑}\langle \{p\} \rangle \text{拦截}\langle \{p\} \rangle)\}$

表 2 列举了该博物馆多路视频监控场景在某夜间时段内解析出的包含人的行为活动(源自其中 13 路相机)的 37 个地理视频镜头,其中,重点表达了视频数据地理关联分析所需要的位置语义和轨迹语义。根据表中数据可知:该组地理视频镜头记录了两名员工和两名来访者的行为,其中来访者 A 和 B 的轨迹语义表明他们分别突破了指定的安全距离,对展品 2 和 1 构成威胁,同时来访者 B 移动了展品 1 的位置。

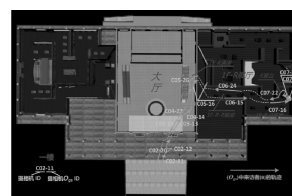
表 2 博物馆多路监控视频实例的三域语义描述
Tab. 2 The three-domain description of geovideo in museum security monitoring

O _{gs} ID	O _{ge}		O _{ob}
	人/展品	地理环境部件	主动行为⟨p⟩/被动行为⟨e⟩轨迹语义⟨c⟩
C02-11	来访者(A) 来访者(B)	大门	⟨A,B⟩向⟨博物馆⟩靠近
C02-12	来访者(B)	大门	⟨B⟩徘徊于⟨大门外⟩
C02-20	来访者(B)	大门	⟨B⟩从⟨博物馆⟩离开
C02-34	来访者(A)	大门	⟨A⟩从⟨博物馆⟩离开
C03-12	员工(D)	大厅 R 楼梯	⟨D⟩来到⟨R 楼梯⟩
C03-13	来访者(A) 来访者(B)	大门 大厅	⟨A,B⟩进入⟨大厅⟩ ⟨A⟩向⟨R 楼梯⟩靠近 ⟨B⟩向⟨1F-R 通道⟩靠近
C03-14	员工(C)	大厅 R 楼梯	⟨C⟩向⟨R 楼梯⟩靠近
C03-18	来访者(B)	大门 大厅	⟨B⟩从⟨博物馆⟩出
C03-29	来访者(A)	大门 大厅	⟨A⟩从⟨博物馆⟩出
C03-31	员工(C) 员工(D)	大门 大厅	⟨C,D⟩从⟨博物馆⟩出
C04-14	来访者(B)	1F-R 通道	⟨B⟩经过⟨1F-R 通道⟩
C04-15	来访者(A)	R 楼梯	⟨A⟩经过⟨R 楼梯⟩
C04-27	来访者(B)	1F-R 通道	⟨B⟩向⟨大厅⟩靠近
C05-16	来访者(B)	1F-R 通道 1F-R 展厅入口	⟨B⟩进入⟨1F-R 展厅⟩
C05-20	员工(C)	1F-R 通道	⟨D⟩经过⟨1F-R 通道⟩
C05-26	来访者(B)	1F-R 通道 1F-R 展厅入口	⟨B⟩从⟨1F-R 展厅⟩出
C06-15	来访者(B)	1F-R-R 展廊	⟨B⟩经过⟨1F-R-R 展廊⟩ ⟨B⟩向⟨1F-R-B 展厅⟩靠近
C06-24	来访者(B)	1F-R-R 展廊	⟨B(X)⟩向⟨1F-R-B 展厅⟩靠近
C07-16	来访者(B)	1F-R-B 展厅	⟨B⟩徘徊于⟨1F-R-B 展厅⟩
C07-17	来访者(B)	K 展台	⟨B⟩停留于⟨K 展台⟩
C07-20	来访者(B)	K 展台	⟨B⟩徘徊于⟨展品 X⟩ (⟨安全距离!⟩)
C07-22	来访者(B) 展品(1)	K 展台	⟨B(X)⟩离开⟨K 展台⟩ ⟨被移动⟨展品(1)⟩,⟨来访者(B)⟩⟩
C08-20	员工(C) 员工(D)	R 楼梯 2F 回廊	⟨C,D⟩从⟨R 楼梯⟩ 到⟨2F⟩

续表 2

O _{gs} ID	O _{ge}		O _{ob}
	人/展品	地理环境部件	主动行为⟨p⟩/被动行为⟨e⟩轨迹语义⟨c⟩
C08-21	来访者(A)	R 楼梯 2F 回廊	⟨A⟩从⟨R 楼梯⟩到⟨2F⟩
C08-25	来访者(A)	R 楼梯 2F 回廊 2F-R 展厅出口	⟨A⟩从⟨2F-R 展厅出口⟩ 到⟨R 楼梯口⟩
C09-11	来访者(A)	2F 回廊 2F-R 展厅入口	⟨A⟩进入⟨2F-R 展厅⟩
C10-14	来访者(A)	展台 L	⟨A⟩徘徊于⟨展台 L⟩
C10-15	来访者(A) 展品 2(2)	展台 L	⟨A⟩靠近⟨展品(2)⟩ (⟨安全距离!⟩)
C10-16	来访者(A)	展台 L 2F-R 展厅出口	⟨A⟩离开⟨展台 L⟩ ⟨A⟩从⟨2F-R 展厅⟩出
C11-33	来访者(A)	2F 回廊	⟨A⟩从⟨R 楼梯口⟩ 到⟨L 楼梯口⟩
C11-35	员工(C) 员工(D)	2F 回廊	⟨C,D⟩从⟨R 楼梯口⟩ 到⟨L 楼梯口⟩
C12-25	来访者(A)	L 楼梯	⟨A⟩向⟨大厅⟩靠近
C12-27	员工(C) 员工(D)	L 楼梯	⟨C,D⟩向⟨大厅⟩靠近
C13-04	员工(D)	1F-L 通道	⟨C⟩向⟨大厅⟩靠近
C13-11	来访者(A)	大厅	⟨A⟩向⟨大门⟩靠近
C13-13	员工(C) 员工(D)	大厅	⟨C,D⟩向⟨大门⟩靠近
C14-18	员工(D)	2F-F 展厅出口	⟨C⟩从⟨2F-F 展厅⟩出

结合表 2 中博物馆室内场景在统一时空框架下的位置语义,该组地理视频镜头中人的行为轨迹如图 2 所示,各地理视频镜头间表达出基于地理实体、语义位置和语义轨迹的关联性,因此有助于实现多尺度事件的理解和推理。本案例中的关联结果显示该时间段多路监控视频数据记录了一个同时包含多阶段和单体引发群体并发性特征的展品(文物)突发安全事件。



(a) 来访者B的关联轨迹



(b) 来访者A的关联轨迹 (c) 员工C、D的关联轨迹

图 2 基于位置语义和轨迹语义的地理视频关联示意
Fig.2 Spatiotemporal association of geovideo based on location semantics and trajectory semantics

由于各地理视频帧、地理视频镜头和镜头组包含了时间、空间、专题(人、展品、威胁事件等)等多维语义信息,因此可灵活支持时空语义一体化的高维索引构建,有助于建立海量监控视频搜索任务的关联约束,显著缩小搜索空间,提高搜索效率。

5 结 论

本文针对传统视频数据模型支持多视频有机关联表示的局限,提出了一种面向特征—行为过程—事件三域的多层次地理视频语义模型,该模型的特点与创新性包括:①通过对地理视频内容变化的表达,综合了地理视频数据和内容变化时空属性的基本维度,实现了视频数据和地理场景的统一描述与映射规则;②突破了传统以摄像机为单位的连续视频流的整体语义描述方法,面向更细化的多粒度层次建立视频语义结构;③将地理环境语义与视频内容语义有机结合,支持多地理视频数据的关联表示。以博物馆公共安全监控的多路视频数据为例进行实例分析,结果表明:模型实现了多地理视频内容和地理环境的语义映射,较好地表达了地理视频面向公共安全事件的数据特征和多视频间的关联性,有助于视频空间全生命周期多尺度危机事件的感知和理解以及建立海量监控视频搜索任务的关联约束,提高地理视频的复杂时空关系的认知计算能力与表达效率。进一步的研究将考虑在模型结构和关系表示的基础上,设计地理视频语义关系的度量方法与计算模型,为分布式地理视频大数据的语义关联的自适应组织、高效管理存储和多约束检索等更深入的应用提供理论支持。

参考文献:

- [1] BERRY J K. Capture “Where” and “When” on Video-based GIS [J]. *GeoWorld*, 2000, 13(9): 26-27.
- [2] KIM S H, ARSLAN AY S, ZIMMERMANN R. Design and Implementation of Geo-tagged Video Search Framework[J]. *Journal of Visual Communication and Image Representation*, 2010, 21(8): 773-786.
- [3] KIM K H, KIM S S, LEE S H, et al. The Interactive Geographic Video [C] // 2003 IEEE International Geoscience and Remote Sensing Symposium: IGARSS'03. Toulouse, France: IEEE, 2003, 1: 59-61.
- [4] CHRISTEL M G, OLLIGSCHLAEGER A M, HUANG Chang. Interactive Maps for a Digital Video Library [J]. *IEEE Multi Media*, 2000, 7(1): 60-67.
- [5] KONG Yunfeng. Design of Geovideo Data Model and Implementation of Web-based VideoGIS [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.)
- [6] NAVARRETE T. Semantic Integration of Thematic Geographic Information in a Multimedia Context [D]. Barcelona, Spain: Universitat Pompeu Fabra, 2006.
- [7] LEWIS P, FOTHERINGHAM S, WINSTANLEY A. Spatial Video and GIS [J]. *International Journal of Geographical Information Science*, 2011, 25(5): 697-716.
- [8] HWANG T H, CHOI K H, JOO I H, et al. MPEG-7 Metadata for Video-based GIS Applications [C] // IEEE International Geoscience and Remote Sensing Symposium: IGARSS'03. Toulouse, France: IEEE, 2003(6): 3641-3643.
- [9] PISSINOU N, RADEV I, MAKKI K. Spatio-temporal Modeling in Video and Multimedia Geographic Information Systems [J]. *GeoInformatica*, 2001, 5(4): 375-409.
- [10] BLOEHDORN S, PETRIDIS K, SAATHOFF C, et al. Semantic Annotation of Images and Videos for Multimedia Analysis [M] // *The Semantic Web: Research and Applications*. Berlin: Springer, 2005: 592-607.
- [11] WANG Xiaofeng, ZHANG Dapeng, WANG Fei, et al. Semantic Trajectory Based Video Event Detection [J]. *Chinese Journal of Computers*, 2010, 33(10): 1845-1858. (王晓峰, 张大鹏, 王绯, 等. 基于语义轨迹的视频事件探测 [J]. *计算机学报*, 2010, 33(10): 1845-1858.)
- [12] AGIUS H W, ANGELIDES M C. Modeling Content for Semantic-level Querying of Multimedia [J]. *Multimedia Tools and Applications*, 2001, 15(1): 5-37.
- [13] LIN C H, LEE A H C, CHEN A L P. A Semantic Model for Video Description and Retrieval [M] // *Advances in Multimedia Information Processing: PCM 2002*. Berlin: Springer, 2002: 183-190.
- [14] AL SAFADI L A E, GETTA J R. Semantic Modeling for Video Content-based Retrieval Systems [C] // 23rd Australasian Computer Science Conference: ACSC 2000. Canberra, ACT: IEEE, 2000: 2-9.
- [15] CHEN Xianming, WANG Xiaoming. The MPEG-7 Video Semantic Description Model Based on Ontology [J]. *Journal of South China Normal University: Natural Science Edition*, 2007(2): 51-56. (陈贤明, 王小铭. 基于本体与 MPEG-7 视频语义描述模型 [J]. *华南师范大学学报: 自然科学版*, 2007(2): 51-56.)
- [16] WANG Yu, ZHOU Lizhu, XING Chunxiao. Video Semantic Models and Their Evaluation Criteria [J]. *Chinese Journal of Computers*, 2007, 30(3): 337-351. (王煜, 周立柱, 邢春晓. 视频语义模型及评价准则 [J]. *计算机学报*, 2007, 30(3): 337-351.)
- [17] KOMPATSIARIS Y, HOBSON P. *Semantic Multimedia and Ontologies [M]*. London, UK: Springer-Verlag Limited, 2008.
- [18] REN W, SINGH S, SINGH M, et al. State-of-the-art on

- Spatio-temporal Information-based Video Retrieval [J]. Pattern Recognition, 2009, 42(2): 267-282.
- [19] LEW M S, SEBE N, DJERABA C, et al. Content-based Multimedia Information Retrieval: State of the Art and Challenges[J]. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP), 2006, 2(1): 1-19.
- [20] ZHU Xingquan, ELMAGARMID A K, XUE Xiangyang, et al. Insight Video: Toward Hierarchical Video Content Organization for Efficient Browsing, Summarization and Retrieval [J]. IEEE Transactions on Multimedia, 2005, 7(4): 648-666.
- [21] ZHU Qing, HU Mingyuan. Lane-oriented 3D Road Network Model[J]. Acta Geodaetica et Cartographica Sinica, 2009, 37(4): 514-520. (朱庆, 胡明远. 基于语义的多细节层次 3 维房产模型[J]. 测绘学报, 2009, 37(4): 514-520.)
- [22] ZHENG Nianbo, LU Feng, LI Qingquan. Dynamic Multi-scale Road Network Data Model for Navigation [J]. Acta Geodaetica et Cartographica Sinica, 2010, 39(4): 428-434. (郑年波, 陆锋, 李清泉. 面向导航的动态多尺度路网数据模型[J]. 测绘学报, 2010, 39(4): 428-434.)
- [23] GONG Jianya, LI Xiaolong, WU Huayi. Spatiotemporal Data Model for Real-time GIS [J]. Acta Geodaetica et Cartographica Sinica, 2014, 43(3): 226-232. (龚健雅, 李小龙, 吴华意. 实时 GIS 时空数据模型[J]. 测绘学报, 2014, 43(3): 226-232.)
- [24] HORNSBY K, EGENHOFER M J. Identity-based Change: A Foundation for Spatio-temporal Knowledge Representation [J]. International Journal of Geographical Information Science, 2000, 14(3): 207-224.

(责任编辑: 宋启凡)

收稿日期: 2014-04-08

修回日期: 2014-06-17

第一作者简介: 谢潇(1986—), 女, 博士生, 研究方向为虚拟地理环境与视频 GIS。

First author: XIE Xiao (1986—), female, PhD candidate, majors in virtual geographic environment and video GIS.

E-mail: xiexiaolmars@gmail.com

通信作者: 朱庆

Corresponding author: ZHU Qing

E-mail: zhuq66@263.net

(上接第 554 页)

- [10] LIU Shanlei, ZHAO Yindi, LI Yingcheng et al. Analysis at Underwater Topographic Two-media Photogrammetry with POS Data [J]. Science of Surveying and Mapping, 2011, 36(6): 42-45. (刘善磊, 赵银娣, 李英成, 等. POS 数据用于双介质水下地形摄影测量的研究 [J]. 测绘科学, 2011, 36(6): 42-45.)
- [11] DANG Yamin, CHENG Pengfei, ZHANG Chuanyin et al. Sea Reefs Mapping Technology and Methods [M]. Beijing: Surveying and Mapping Press, 2012. (党亚民, 程鹏飞, 章传银等. 海岛礁测绘技术与方法 [M]. 北京: 测绘出版社, 2012.)
- [12] QUAN X, FRYE S. Empirical Equation for the Index of Refraction of Seawater [J]. Applied Optics, 1995, 34(18).
- [13] JIANG Lianguang, SUN Peng. Models of the Wavelength Dependence for the Index of Refraction of Water [J]. Chinese Journal of Spectroscopy Laboratory, 2002, 19(4): 554-556. (姜良广, 孙鹏. 水的折射率与波长间函数关系的模型 [J]. 光谱实验室, 2002, 19(4): 554-556.)
- [14] Lianyungang Maritime Safety Administration. (2011-03-15) [2014-01-30]. <http://www.lygmsa.gov.cn/>. (连云港海事局网站. (2011-03-15) [2014-01-30]. <http://www.lygmsa.gov.cn/>.)
- [15] ZHANG Zuxun, ZHANG Jianqing. Digital Photogrammetry [M]. Wuhan: Wuhan University Press, 2012. (张祖勋, 张剑清. 数字摄影测量 [M]. 武汉: 武汉大学出版社, 2012.)
- [16] SHEN Jiashuang. Research on Technology of Equal Waterline Information Extraction and Vertical Datum Transformation in Coastal Zone [D]. Zhengzhou: Information Engineering University, 2011. (申家双. 海岸带等水位线信息提取与垂直基准转换技术研究 [D]. 郑州: 信息工程大学, 2011.)
- [17] BAO Jingyang, ZHANG Chuanyin. On the Sea and Ocean Vertical Datum [J]. Bulletin of Surveying and Mapping, 2010, (6): 10-11, 26. (暴景阳, 章传银. 关于海洋垂直基准的讨论 [J]. 测绘通报, 2010(6): 10-11, 26.)
- [18] LI Tao. Construction and Visualization Methods and Implementation for Delaunay Triangulation [D]. Nanchang: East China Institute of Technology, 2012. (李涛. Delaunay 三角网构建及可视化方法与实现 [D]. 南昌: 东华理工大学, 2012.)
- [19] XU Duowen. Construction and Application of TIN [D]. Ganzhou: Jianxi University of Science and Technology, 2010. (许多文. 不规则三角网 (TIN) 的构建及应用 [D]. 赣州: 江西理工大学, 2010.)
- [20] SHEN Jiashuang, ZHAI Jingsheng, ZHAI Guojun et al. The Research on the Coastal Topographic Map and Its Surveying Method [J]. Bulletin of Surveying and Mapping, 2007(8): 29-32. (申家双, 翟京生, 翟国君, 等. 海岸带地形图及其测量方法研究 [J]. 测绘通报, 2007(8): 29-32.)

(责任编辑: 张艳玲)

收稿日期: 2014-05-26

修回日期: 2014-11-14

第一作者简介: 周高伟(1988—), 男, 硕士生, 研究方向为航空航天双介质摄影测量。

First author: ZHOU Gaowei (1988—), male, post graduate, majors in two-media photogrammetry of aeronautics & astronautics.