

基于视频的地理场景增强表达研究

宋宏权,刘学军,闫国年,王美珍

(南京师范大学虚拟地理环境教育部重点实验室,江苏 南京 210046;南京师范大学地理科学学院,江苏 南京 210046)

摘要:视频具有表达直观、信息丰富、动态实时等特点,将其与 GIS 有机集成形成地理视频(GeoVideo),可更加形象的表达地理空间。该文利用视频分割技术将视频分割为具有不同摄像机运动状态的视频片段,针对摄像机运动的视频片段,采用摄像机跟踪(Camera Tracking)技术求算视频帧对应的摄像机内外参数,并基于地理视频数据模型,采用 KML/XML 描述视频帧、摄像机位置、摄像机姿态等之间的时空映射关系。最后,设计了原型系统,在网络环境中聚合地理信息服务、视频流媒体服务与 KML 文档服务等,实现了视频与三维地理场景的交互与融合。实验表明,基于该地理视频数据模型,利用视频数据可实现对地理场景的增强表达,使用户能更形象的认知地理空间。

关键词:地理视频;视频 GIS;摄像机跟踪;地理空间表达

中图分类号:P208 文献标识码:A 文章编号:1672-0504(2012)05-0006-04

0 引言

地理空间表达一直是地理信息科学领域研究的重要内容,也是地理信息科学研究面临的一个重大挑战。地理空间表达是一个空间认知、信息转换与信息传输的交互过程^[1],经历了由自然语言、地图到 GIS 的演变过程^[1,2]。由于现实地理世界具有复杂多样、三维立体、时空动态等特性,目前对地理空间的表达依然是抽象的二维地图占主导地位,为解决语义差问题,出现了虚拟地理环境、地理增强现实、地理超媒体等多种地理空间表达方式。

地理视频(GeoVideo)是集成地理信息与视频片段进行地理空间表达的新方式,国内外学者对视频与 GIS 的集成进行了理论和应用探索。1978 年麻省理工学院的 Lippman 教授,首次将视频与空间数据集成,开发了动态、交互式超媒体地图^[3]。随后多媒体技术逐步引入 GIS 领域,提出了多媒体地图、多媒体 GIS、地理超媒体等概念。近年来,国外一些学者将视频和地理信息相结合开展了相关研究:提出了视频地图框架,并设计了数据的外业采集、处理与应用的概念方案^[4];建立了视频片段地理索引,实现了相关的数据采集和应用系统^[5-8];对视频影像中的地物实体进行时空建模并构建视频元数据,完成了空间位置与视频帧的映射^[9,10]。同时,我国学者也进行了相关研究,如将视频与地理信息集成分别应用于铁路、公路的可视化管理^[11-13],设计基于地

图的车载移动视频监控系系统^[14]以及便携式可定位视频系统的研发^[15]、地理视频数据模型设计及网络环境下地理视频的应用^[16-20]等。

总体看,目前对视频与 GIS 的集成研究是将视频作为空间数据的属性,或将其与 GPS 数据融合实现视频与二维电子地图的交互,但将视频与三维 GIS 的交互、融合研究较少,不能有效利用视频本身丰富的信息。本文尝试利用视频分割、摄像机跟踪(Camera Tracking)、聚合(Mashup)等技术在网络环境下进行视频与三维地理场景间的交互与融合,以实现基于视频的地理场景增强表达并验证其可行性。

1 数据模型与数据组织

1.1 地理视频数据模型

本文基于文献^[16]的地理视频数据模型,将其扩展至三维,其实体—关系如图 1 所示。视频具有多种格式和标准,但均可抽象为视频片段与视频帧,视频片段由多个连续的视频帧组成。视频片段具有对应的摄像机轨迹,视频帧图像具有对应的摄像机位置与姿态,为视频片段添加元数据描述可用于视频片段的语义检索,对视频帧图像中的地理实体进行描述可实现视频帧的语义检索。利用文本、KML/XML、关系数据库等,描述摄像机轨迹、摄像机位置与姿态、三维地理场景的镜头位置与姿态及三维地理实体间的关系,建立视频与三维地理场景之间的时空/语义映射关系,将视频的真实表达与三

收稿日期:2012-04-23; 修订日期:2012-06-20

基金项目:国家科技支撑计划项目“视频 GIS 与突发公共事件的感知控制系统”(2012BAH35B02);江苏省高校自然科学重大基础研究资助项目(10KJA420025)

作者简介:宋宏权(1986-),男,博士研究生,主要从事视频 GIS、GIS 应用研究。E-mail:hongquansong@126.com

维地理场景的抽象表达相结合,以期实现基于视频的地理场景增强表达,从而在一定程度上克服传统 GIS 的抽象与静态表达特性。

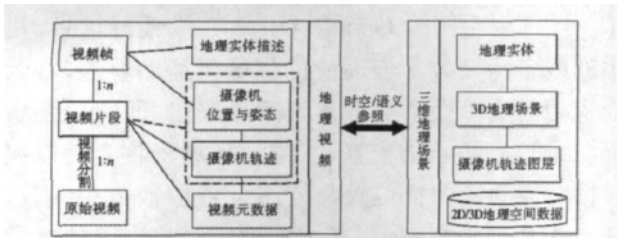


图 1 地理视频数据模型

Fig. 1 Data model of GeoVideo with 3D geographical scene

1.2 数据处理与组织

基于地理视频数据模型,对采集的视频数据和 GPS 数据进行处理。首先,通过时间信息对 GPS 数据和视频数据进行配准,根据摄像机运动和空间位置信息进行视频分割,即将视频序列按一定标准分割成多个视频片段^[21]。摄像机拍摄视频时具有平移、缩放、旋转及静止等不同的运动状态,为恢复摄像机的内外参数,不同运动状态的视频片段具有不同的处理方式,故将视频根据摄像机的运动状态分割成不同类型的视频片段,同时将 GPS 数据分割成与视频片段对应的轨迹片段。

摄像机跟踪技术是利用计算机视觉领域的运动恢复结构(Structure from Motion, SFM)等技术,通过对摄像机拍摄的图像序列或视频进行特征点检测、特征点匹配、摄像机自标定等处理,分析视频或图像序列中物体的运动,完成摄像机内外参数的求解及场景三维结构的恢复。由于摄像机跟踪技术无法处理静止状态的视频片段,故本文只处理具有平移运动或同时具有平移、缩放及旋转运动的视频片段。本文利用自动摄像机跟踪软件(ACTS)^[22]对视频片段进行摄像机参数求解。对视频片段进行摄像机跟踪处理后,可得到摄像机在整个运动过程中的姿态变换参数以及摄像机坐标系下的摄像机运动轨迹(图 2,见封 3),其中,左侧为特征点检测与匹配,右侧为求得的摄像机运动轨迹与姿态变换参数。

摄像机跟踪求解完成后,基于地理视频数据模型,利用 KML 文档描述摄像机姿态、位置、轨迹及三维场景中虚拟镜头位置与姿态之间的关系。在 KML 2.2 版本中,增加了新元素 <Camera>,它提供了一种能够指定虚拟镜头(观察者视点)与相关视图参数的方法来查看 3D 场景。本文简要说明如何利用 KML 描述各参数间的关系,例如:

```
<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://www.opengis.net/kml/2.2" xmlns:
```

```
gx="http://www.google.com/kml/ext/2.2">
<Document><name>视频片段 1.flv</name>
<gx:Tour><gx:Playlist><gx:FlyTo>
<gx:duration>1.0</gx:duration>
<Camera>
<longitude>114.3033817</longitude>
<latitude>34.82331667</latitude>
<altitude>1.5</altitude>
<heading>179.904546</heading>
<tilt>45.00884</tilt>
<roll>-0.032959</roll>
</Camera>
</gx:FlyTo>...
</gx:Playlist></gx:Tour>
<Folder>
<Placemark><name>摄像机轨迹</name>
<LineString><coordinates>
114.3034,34.8234,0 114.3033,34.8234,0...</coordinates>
</LineString>
</Placemark></Folder></Document></kml>
```

如图 3, KML 文档的 <Camera> 中元素的意义分别如下:<longitude>为虚拟镜头的经度;<latitude>为虚拟镜头的纬度;<altitude>为镜头距地表的距离(m);<heading>为镜头绕“Z”轴旋转的角度(方位角,°,其范围为 0°~360°;<tilt>为镜头绕“X”轴旋转的度数,0°表示视线垂直俯瞰场景,90°表示视线朝向地平线,大于 90°表示视线上指天空;<roll>为镜头绕“Z”轴的第二次旋转,其旋转变换顺序为,首先沿“Z”轴转换到 <altitude>高度,然后利用 <heading>绕“Z”轴旋转,接着按 <tilt>值绕“X”轴旋转,最后利用 <roll>值绕“Z”轴第二次旋转,从而完成镜头的变换操作。

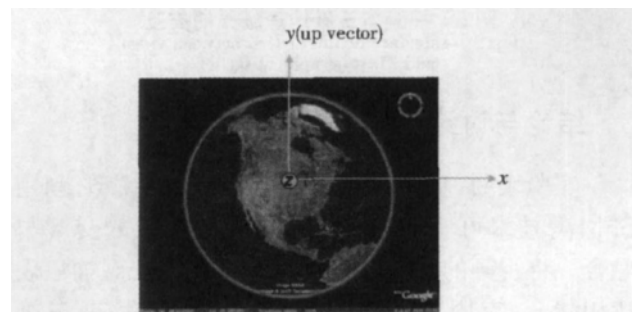


图 3 三维地理场景中虚拟镜头示意(引自 Google 网站)
Fig. 3 Virtual camera in 3D geographical scene (cited from Google website)

2 系统设计与实现

2.1 系统流程设计

本系统包括数据采集、数据处理、数据服务与应用 4 部分(图 4)。具体为:1)利用摄像机和 GPS 接

收机分别采集视频和 GPS 数据;2)对原始视频进行视频分割、SFM 摄像机跟踪等,基于本文的地理视频数据模型,利用 KML 描述摄像机位置、轨迹、姿态以及三维地理场景中虚拟镜头位置与姿态等参数间的时空映射关系;3)基于 Web 服务发布视频片段、KML 描述文档;4)利用 Mashup 技术聚合三维地理空间数据服务、视频流媒体服务及 KML 文档服务,从而实现网络环境下基于视频的三维地理场景增强表达。其中,Mashup 技术是将互联网上的不同资源整合到一起实现新的应用。

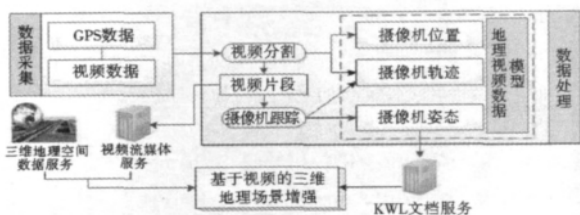


图 4 系统流程设计
Fig. 4 System process design

2.2 系统实现

通过数据采集、数据处理等操作,将视频片段、KML 文档等以 Web 服务方式发布,利用 JavaScript 脚本语言,结合 Google Earth API 开发专用地理视



图 5 视频与三维地理场景的交互
Fig. 5 Interface of interaction between video and 3D geographical scene

3 结论与讨论

实验表明,基于本文的地理视频数据模型,利用 SFM 等技术可实现视频与三维地理场景的交互与融合。将视频的微观真实表达与 GIS 的宏观抽象表达相结合,克服了传统二维 GIS 的抽象性及三维 GIS 的虚拟性,使空间数据的抽象性与视频数据的直观、形象等特点相互补充。空间数据对视频的补充,使人们更好地理解视频数据的空间特征;视频对空间数据的补充,更易于人们利用视频认知对应的真实地理空间,身临其境地认知地理环境,充分体现了侧面看世界与高空看世界相结合的优越性。

频播放器^[19,20],聚合 Google Earth 三维空间数据服务、视频流媒体服务与 KML 文档服务,最后实现基于视频的三维地理场景增强表达,如图 5 和图 6 所示。图 5 中左侧为 Google Earth 三维地理场景,其中的线条为视频片段对应的摄像机运动轨迹,右侧为视频播放界面。点击视频播放按钮,即实现视频与三维地理场景的交互,三维地理场景将实时以视频帧对应的摄像机位置与姿态设置虚拟镜头参数。任意拖放选择视频帧,三维地理场景将以该帧对应的摄像机位置与姿态设置虚拟镜头参数;同理,选择左侧摄像机轨迹时间轴某位置,三维地理场景以此位置对应的摄像机姿态设置虚拟镜头,视频播放器播放该位置对应的视频帧,即实现了视频帧与三维地理场景的交互索引查询操作。

根据视频帧对应的摄像机位置与姿态,在三维地理场景中模拟摄像机成像过程,即可将视频图像在三维地理场景中以其对应的摄像机姿态融合显示,实现图像与三维地理场景的融合。视频播放过程中,点击融合查看按钮(图 6),可重现当前播放视频帧在三维地理场景中的位置与姿态,使人们更直观的认知地理环境。



图 6 视频与三维地理场景的融合
Fig. 6 Interface of fusion between video and 3D geographical scene

本系统只是基于 Google Earth 验证了数据模型的可行性,对其进一步扩充,可应用于突发公共事件管理、应急指挥、智能交通、设施管理、旅游宣传等领域。视频 GIS 还是一个新的研究领域,本文只实现了网络环境下视频与三维地理场景的交互与融合表达,未充分利用视频中丰富的信息,在可量测视频三维场景重构、视频与多源地理数据的集成与表达、地理增强现实、基于视频的地理空间分析等方面都需要更深入的研究。

参考文献:

[1] 韩志刚,孔云峰,秦耀辰.地理表达研究进展[J].地理科学进

- 展,2011,30(2):141-148.
- [2] 林琿,黄凤茹,鲁学军,等.虚拟地理环境认知与表达研究初步[J].遥感学报,2010,14(4):822-838.
- [3] LIPPMAN A. Movie maps: An application of the optical videodisc to computer graphics[J]. SIGGRAPH'80,1980,14(3):32-43.
- [4] BERRY J K. Capture 'Where' and 'When' on video-based GIS[J]. GEOWORLD,2000,13(9):26-27.
- [5] LEE S Y, KIM S B, CHOI J H. 4S-Van: A prototype mobile mapping system for GIS[J]. Korean Journal of Remote Sensing, 2003,19(1):91-97.
- [6] HWANG T H, CHOI K H, JOO I H, et al. MPEG-7 metadata for video-based GIS applications[A]. IGARSS'03 Proceedings [C]. 2003. 3641-3643.
- [7] LEE S Y, CHOI K H, JOO I H, et al. Design and implementation of 4S-Van: A mobile mapping system[J]. ETRI Journal, 2006,28(3):265-275.
- [8] NAVARRETE T, BLAT J. VideoGIS: Segmenting and indexing video based on geographic information[A]. 5th AGILE Conference on Geographic Information Science[C]. Palma de Mallorca, Spain, 2002. 1-9.
- [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] JOO I H, HWANG T H, CHOI K H. Generation of video metadata supporting Video-GIS integration[A]. ICIP'04 Proceedings[C]. Singapore, 2004. 1695-1698.
- [11] 唐冰,周美玉.基于视频图像的既有线路地理信息系统[J].铁路计算机应用,2001,10(11):31-33.
- [12] 孔云峰.一个公路视频GIS的设计与实现[J].公路,2007(1):119-121.
- [13] 李郁峰,朱金陵.铁路线路视频数据采集系统设计与开发[J].铁路计算机应用,2004,13(12):4-6.
- [14] 丰江帆,张宏,沙月进. GPS车载移动视频监控系统的的设计[J].测绘通报,2007(2):52-54.
- [15] 吴勇,刘学军,赵华,等.可定位视频采集方法研究[J].测绘通报,2010(1):24-27.
- [16] 孔云峰.地理视频数据模型及其应用开发研究[J].地理与地理信息科学,2009,25(5):12-16.
- [17] 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.
- [18] 孔云峰.基于Web服务的地理超媒体系统设计开发与应用[J].地球信息科学学报,2010,12(1):76-82.
- [19] 宋宏权,孔云峰. Adobe Flex 框架中的视频GIS系统设计与开发[J].武汉大学学报(信息科学版),2010,35(6):743-746.
- [20] 宋宏权,孔云峰. Flex 框架下网络视频GIS设计与实现[J].测绘科学,2010,35(5):208-210.
- [21] 任菲,刘学军,丰江帆,等.基于空间信息辅助的视频分割研究[J].计算机应用研究,2009,26(4):1546-1548.
- [22] DONG Z, ZHANG G, JIA J, et al. Keyframe-based real-time camera tracking[A]. 2009 IEEE International Conference on Computer Vision (ICCV) [C]. Kyoto, Japan, 2009. 1538-1545.

Research on Augmenting Aerial Earth Maps Using Videos

SONG Hong-quan, LIU Xue-jun, LV Guo-nian, WANG Mei-zhen

(Key Laboratory of Virtual Geographic Environment, Ministry of Education, Nanjing Normal University, Nanjing 210046; School of Geography Science, Nanjing Normal University, Nanjing 210046, China)

Abstract: Video clip can provide rich information by the means of intuitive expression, real-time expression, and so on. GeoVideo, the video with geographic description, is a new approach to the geospatial representation. This paper designs a GeoVideo data model of interactive geospatial expression between video and 3D geographical scene. Video files were divided into video clips by different camera motions using video segmentation technique. The camera parameters of video frames were estimated through camera tracking technique. Then KML documents were used to describe mapping relationships among video frames, camera positions and camera poses based on this GeoVideo data model. Finally, a web service-based prototype system was developed. These types of web services, i. e. 3D web mapping service, web video service and KML documents service, were 'mashuped' using google earth API. Users can experience the interaction between 3D geographical scene and video, this overcomes the traditional abstraction of 2D GIS and virtuality of 3D GIS, so people can more intuitive understand the geographic environment. Case study shows that the data model in this paper is feasible for augmenting aerial earth maps using videos, and this data model has application potentials in facility management, tourism promotion and other fields.

Key words: GeoVideo; video GIS; camera tracking; geospatial representation

第 7 页插图

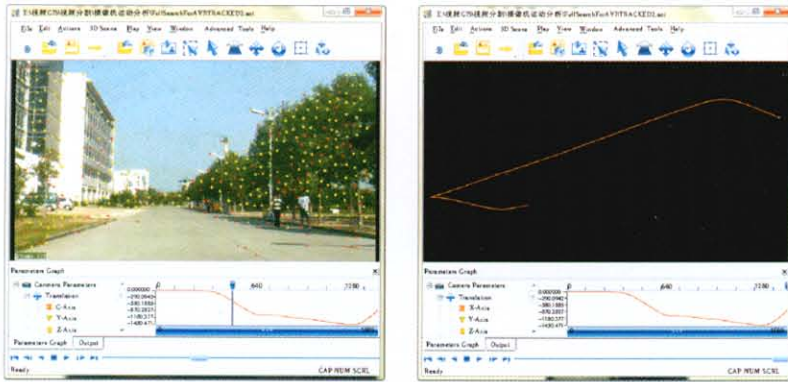


图 2 摄像机跟踪
Fig. 2 Camera tracking results

第 81 页插图

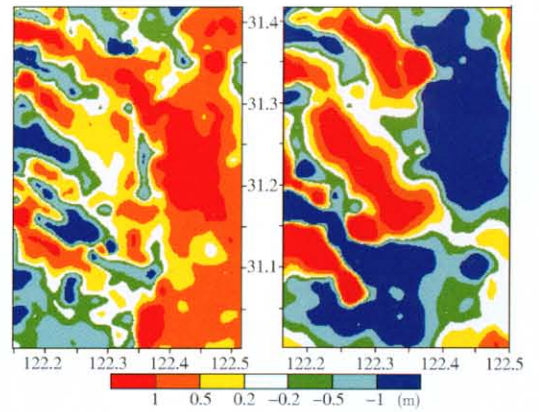


图 2 研究区冲淤厚度分布
Fig. 2 Spatial distribution of accretion-erosion thickness in study area

第 37 页插图

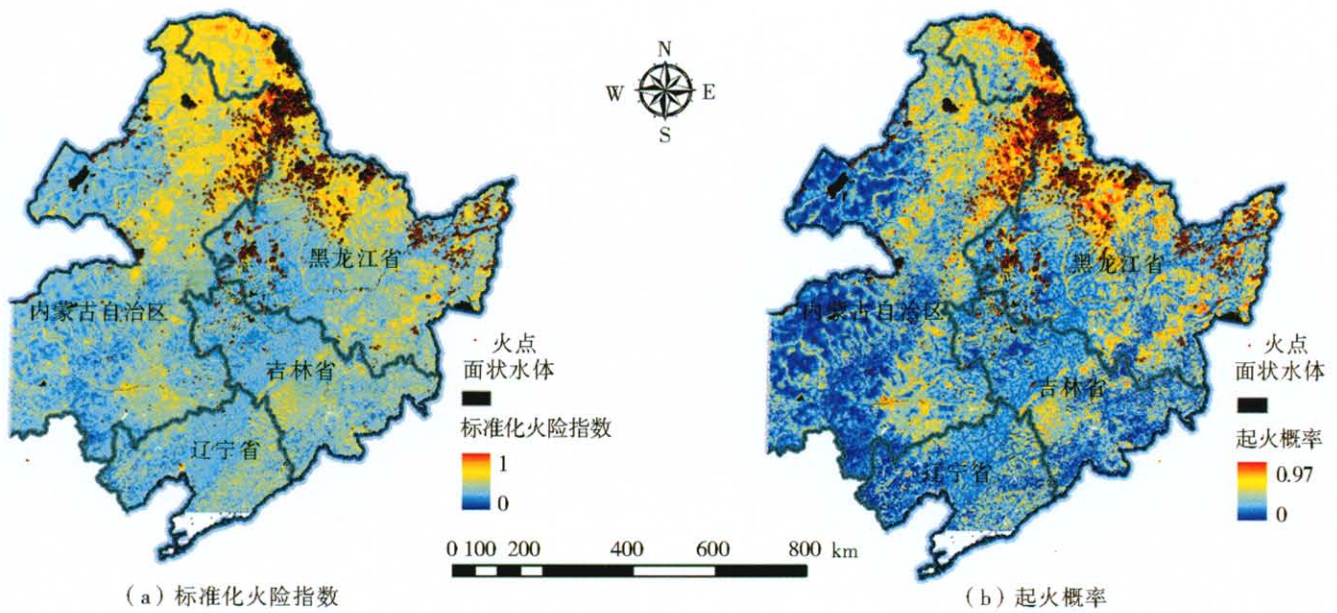


图 1 标准化火险指数和起火概率

Fig. 1 The standardized fire danger index map from frequency ratio and fire occurrence probability map from binary logistic regression

第 58 页插图

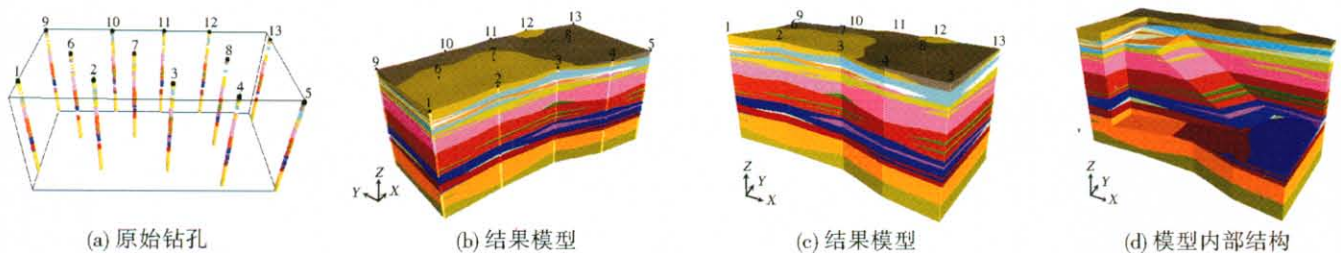


图 5 某小区建模实例
Fig. 5 A practical case