



Geographic Video 3D Data Model Design and Application

Chen Zhilong¹, Wang Xingbin², Yang Hongyu³ and Yang Xiaobin⁴

The Research Center on Underground Space, ^bPLA University of Science & Technology, Nanjing, China

ABSTRACT

Video is valuable for true geographic representation because of the spatial-temporal information of it. With the popularity of video capture devices such as smartphones, the volume of user-generated geographic video clips has grown significantly and the trend of this growth is quickly accelerating. Such a massive and increasing volume poses a major challenge to efficient video management. This paper aimed to introduce a geographic video 3D data model based on geographic information. The main idea of the model is to utilize the location, trajectory and azimuth information acquired by sensors such as GPS receivers and 3D electronic compasses in conjunction with video contents. The raw spatial information is synthesized to point, line, polygon and solid according to the camcorder parameters such as focal length and angle of view. With the video segment and video frame, the three categories geometry object was defined using the geometry model of OGC Simple Features Specification for SQL. Additionally, the application case of the model including video query, video measure and video playing is discussed in detail. The application indicates that the model is an integration, loosely coupled, flexible and extensible data model for the management of geographic video.

Keywords: Geographic Video, Data Model, 3D, GIS

INTRODUCTION

As a normal medium, video contains a rich set of spatial and attribute information. The geospatial that video represent has a sense of reality and a wealth of detailed information. Video and geographic information can be integrated, that is geographic video, and it includes both spatial and temporal geographic features acquired through ground-based or non-ground-based cameras (Furht B., 2008). Geographic video is a multimedia contains location information. The geo-tag is attached to the video clip through the mapping of stereo video frame and location information. It supports location-based video indexing and query. Geographic video can fully play the advantage of video and GIS. It overcomes some of the disadvantages, such as non-intuitivism of 2D GIS, the modeling complexity of 3D GIS, the getting lost in space of video. Based on the acquirement convenience and modeling or analysis flexibility of video data, geographic video can provide a perspective and agile way to observe the geographic world more naturally, visually and timely, and implement image measure.

With the popularity of video capture devices, such as smartphones, the volume of user-generated video clips has grown significantly. This trend of growth is quickly accelerating; e.g., Internet users are uploading 100 hours of video per minute to YouTube. Such a massive and increasing volume poses a major challenge to efficient video management with geographic information. But the geographic video data model is discussed much little.

In this paper, we proposed a geographic video 3D data model (3DGV for short) for video data management. Our framework stressed the following key issues: (1) video with geographic information data analysis and a data model; (2) 3DGV application case. The goal is to enhance video management using geographic information. The remainder of this paper is organized as follows. Section 2 analyzes the state-of-the-art studies and technology of geographic

video. Section 3 details the 3DGV based on data analysis. Section 4 discusses the data acquisition of 3DGV with geographic information. The application case including video query and playing is shown and explained in Section 5. Finally, Section 6 concludes the paper.

RELATED WORKS

There are rich achievements in geographic video in recent years. The Redhen Company introduced the image multimedia mapping system in 1998. The system collects the geographic video data by GPS receiver and camcorder; produce the geographic video database by GIS. Berry proposed the video map system framework in 2000 (Berry, 2000). In VideoGIS project, Navarrete and Blat integrated the video and geographic information, built the spatial index of video clip, make the hypervideo invoked in geographic context (T. Navarrete and J. Blat, 2002). Diego L. et al. implemented a shallow-water positioning system (SWaPS) based on geographic video, and used to monitor the benthic habitats (D. Lirman, G. Deangelo et.al., 2008). Edmundo J. B. developed a YonaPMS.Video system to manage the road asset and evaluate or compare the road condition through the integration of high resolution digital video and GIS. YonaPMS.Video has been used in Israel and other countries in numerous projects (E. Botner and M. Hoffman, 2009).

In the InfoMedia project, Michael G. Christel utilized the location information in news stories to query the video library. Christel et al. recognized and extracted the location narrative with natural language processing and context adjustment methods. These location narratives were matched with a place name dictionary for geocoding and added to news video as geographic information metadata. Based on these metadata, the two-way connection between place names and news video was built. One can quickly find the news video clips according to place names, highlight the place in the text and map when video was played. Because of the geographic information, one can quickly query news videos (Christel, Olligschlaeger, & Huang, 2000). Xiaotao Liu designed and implemented the automatic video annotation and querying system, named SEVA, with rich sensor information. The SEVA system filtered and refined the query results through the adjacent and location information recorded in video streams (Liu, Corner, & Shenoy, 2005). Sakire Arslan Ay et al. extensively studied the video retrieval method based on location information. (1) Arslan et al. proposed a video retrieval method based on camera position. In this method, the GPS receiver and digital compass were integrated with the camera. The position, orientation and field of view of the video clips were collected automatically. The viewable scene model based on this sensor information was designed to index or query the video. A prototype system with data acquisition, storage and video retrieval functions had been implemented to argue the advantage of this method. (2) Arslan et al. realized the geo-referenced video search engine on the web. The user can query geo-referenced video according to the area of interest. The map user interface helped users better understand video on one location. (3) Arslan et al. introduced the vector-based field of view model. This model can be used for different video retrieval methods, such as a given point search, distance constraint point retrieval, the direction constraint point search, distance and direction constraints point retrieval and rectangular area retrieval. Query experiments for all sorts of methods were fulfilled. The results show that the model can effectively improve video retrieval precision and search efficiency. (4) Arslan et al. discussed the relevance-ranking algorithm and grid-based index in geo-referenced video searches (Ay, 2010; Kim, Ay, Yu, & Zimmermann, 2010; Ay, Zhang, Kim, He, & Zimmermann, 2009; Ay, Zimmermann, & Kim, 2008, 2010; Ma, Ay, Zimmermann, & Kim, 2013). The Open Geospatial Consortium (OGC) defined the view cone model for the video frame. It is a closed polygon describing either the camera view cone when looking over the horizon, or the quadrangle that shows the area of interest when looking down at the ground. The view cone polygon is composed of three or five points. Furthermore, the storage string format for the view cone polygon was defined, and the geo-video web service was implemented based on simple object access protocol (SOAP) (J. Lewis, 2006).

Paul Lewis detailed the link and integration of spatial video and geographic information systems. The spatial video is a specific extension to any video formats where spatial attributes have been applied to the frames within the video sequence. These spatial attributes include frame location, orientation and trajectory. A data structure called viewpoint was proposed. It is an extension of the OGC view cone model and represents the capture location and geographical extent of each video frame. The camera calibration model, camera geometric equations and parameter calculation method were also discussed in detail with a 2d viewpoint model as an example. The 2d viewpoint spatial database was designed with seven relational table schemas, and the interpolation, matching and drift of GPS location was analyzed. Finally, the video retrieval methods based on 2d viewpoint database were discussed, and the spatial video player was developed (P. Lewis, 2009; P. Lewis, Fotheringham, & Winstanley, 2011).

GEOGRAPHIC VIDEO 3D DATA MODEL VIDEO DATA ANALYSIS

Video data is more complex and larger in volume than the traditional data. They usually combine visual and audio data, as well as textual data. The traditional video data model is a hierarchical structure which includes five levels with video frame, video shot, video scene, video segment and video sequence (Amr Ahmed, 2009). The frame is a

single still image. It is the smallest logical unit in this hierarchy. The shot is a sequence of frames that has been recorded continuously. The scene is a collection of related shots. The video segment is a group of scenes related to a specific context. A video sequence consists of a number of video segments. There is the composed-of relation between the five levels (Figure 1).

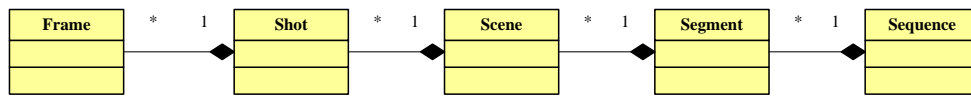


Fig.1 Video Objects Graph

Because the geographic video is recorded along the road or street using the mobile platform, we redefined the traditional video data model for the sake of the simplification in this paper. We defined the two levels of the video object. One is video frame, another is video segment. The video frame is the still image which was recorded in a position of the path at the specified time. The video segment is the collection of the frames between the starting point and the end point of the path which was recorded spatial-temporal continuously.

CONCEPTUAL MODEL

With the video segment and video frame, we defined the three categories geometry object using the geometry model of OGC Simple Features Specification for SQL (Figure 2). It include as follows.

(1) Video location

We can record the location and attitude information when videos are shooting. Based on this information, we defined the VFLocation and VSLocation object to present the video frame and video segment location. These two objects are inherited from point object of OGC Geometry model. Its form is in ①.

$$(x, y, z, yaw, pitch, roll) \quad \textcircled{1}$$

Where x, y, z is the video shooting coordinates, yaw, pitch, roll denote the camera attitude.

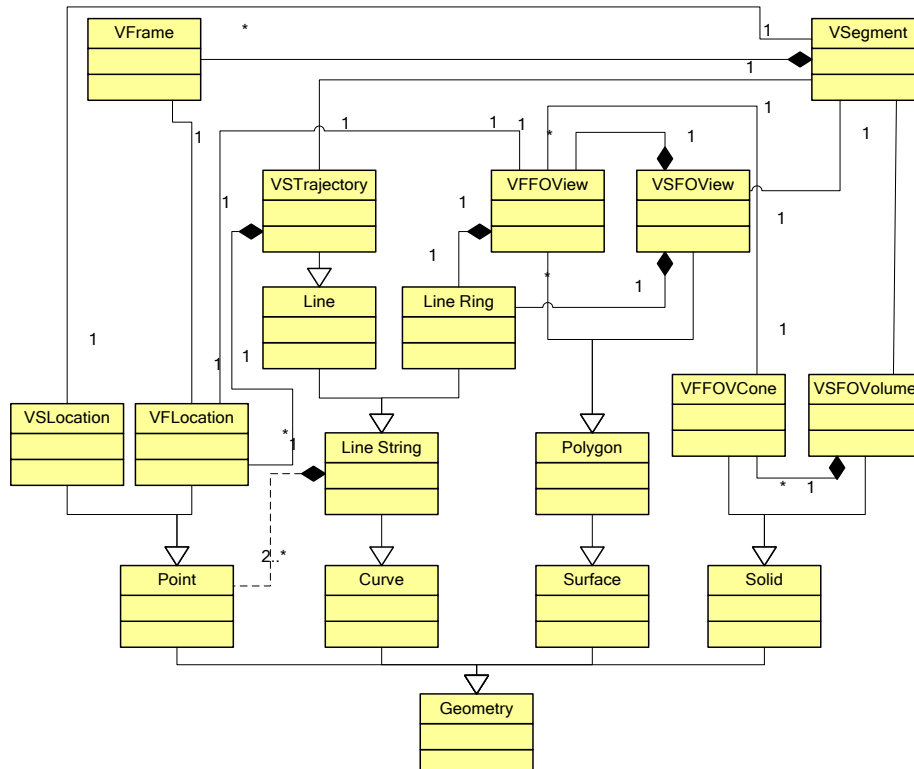


Fig. 2 UML View of 3DGV

(2) Video trajectory

We can generate the shooting trajectory for video segment using the video frame location if the vehicle-mounted or flight-mounted carriers are used to shoot video. We defined VSTrajectory object to present video segment trajectory

which is inherited from line object of OGC Geometry model. Its form is in ②.

$$\{VFLocation_i\}, i=1,2,3\dots \quad \textcircled{2}$$

Where: i denote the video frame sequence number.

(3) Video field of view (FOV) in 2D

Each video segment or frame corresponds to the real world scene. We can generate a 2D fan shape for video frame FOV using the VFLocation and other parameters such as max visible distance theoretically. This fan can be defined by: vertex, the shooting position; radius, the max visible distance; azimuth, the shooting yaw; angular, the camera horizontal field of view (Figure 3). We abstracted the VFFOVView to present the video frame FOV which is inherited from polygon object of OGC Geometry model. Its form is in ③.

$$\{x_i,y_i\}, i=1,2,3\dots \quad \textcircled{3}$$

Where: xi, yi is the coordinates of the 2D polygon.

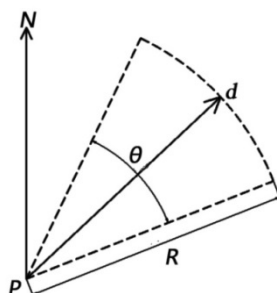


Fig.3 Theoretical field of view of Frame in 2D

For Video segment, all VFFOVViews for each video frame can be union to one irregular polygon. Thus the VSFOView object is defined to present the video segment field of view. Its form is in ④.

$$\cup(VFFovView_i) \quad \textcircled{4}$$

(4) Video field of view (FOV) in 3D

We can also generate a solid for video frame FOV in 3D space. This 3D FOV is a view frustum towards the shot target in a non-occluded environment theoretically. Like 2D FOV, This view frustum can be defined by four elements. But the angular are include horizontal and vertical field of view. We defined VFFovCone object to present video frame 3D FOV (Figure 4). It is inherited from solid object of OGC Geometry model. Its form is in ⑤.

$$\{x_i,y_i,z_i\}, i=1,2,3\dots \quad \textcircled{5}$$

Where: xi, yi ,zi is the point coordinates of the 3D Solid.

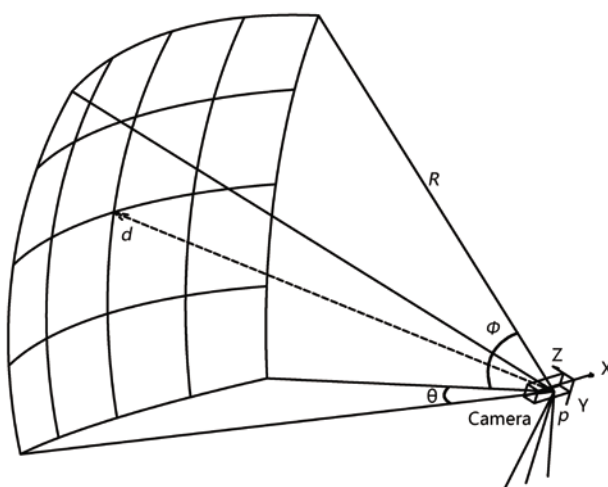


Fig.4 Theoretical field of view of Frame in 3D

For Video segment, all VFFOVCones for each video frame can be union to one irregular Solid. Thus the

VSFOVCone object is defined to present the video segment 3D field of view. Its form is in ⑥.

$$U(VFFovCone_i) \textcircled{6}$$

DATABASE REPRESENTATION

As the database representation of 3DGV, there are 9 feature tables created to represent the logical structures for integrated management of geographic video in 2D/3D space. The relationships within the table are illustrated in figure 5.

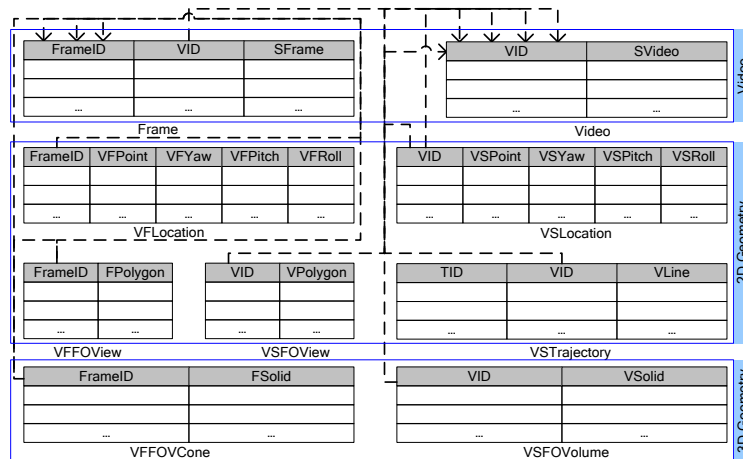


Fig.5 Database representation of 3DGV

Video tables include the frame table and the video table, which are linked by VID property. The primary key of the frame table is FrameID, which is used to link the geometry tables. The video and frame data are stored in the SVideo and SFrame property. The form of video or frame data can be the data path with string type or the LOB type in database.

Geometry tables include the VFLocation/VSLocation table, the VSTrajectory table, the VFFOVView/VSFOView table and the VFFOVVolume/VSFOVolume table, which are used to store spatial information for the video and frame. The primary key of the VFLocation, VFFOVView and VFFOVVolume table is FrameID, which refers to the Frame table primary key. The primary key of the VSLocation, VSTrajectory, VSFOView and VSFOVolume table is VID, which refers to the Video table. The spatial data is stored in VFPoint, VSPPoint, FPolygon, VPolygon, VLine, FSolid or VSolid with related forms, such as Oracle SDO_Geometry. The camera attitude data is stored in FYaw, FPitch, and FRoll property.

CONCLUSION

In this study, the geographic video 3D data model and the application case with it was introduced. The model was composed of entities of video and geometry. The model objects and their relations were defined in Unified Model Language carefully. At logic levels, 9 core relation tables were designed for application purpose. The data acquisition of the data model was detailed, and then used the GPS receiver to capture the position, and the 3D digital compass to collect the attitude of the camera. We took an example of the field of view to illustrate the calculation of the theoretical and actual FOV. The application case shows that many of the fundamental aspects of our proposed data model can be effectively instantiated. Further development of this model is about the design and implementation of application system as well as the efficiency evaluation.

REFERENCES

[1]Amr Ahmed(2009). Video Representation and Processing for Multimedia Data Mining, in Semantic Mining Technologies for Multimedia Databases, Dacheng Tao, Dong Xu, and Xuelong Li Editors. IGI Global: New York. pp.1-31.
 [2]Arslan Ay, S. (2010). Leveraging georeferenced meta-data for the management of large video collections.University Of Southern California.
 [3]Arslan Ay, S., Zhang, L., Kim, S. H., He, M., & Zimmermann, R. (2009). GRVS: A Georeferenced Video Search Engine. Paper presented at the ACM MM09, pp. 977-978.

[4]Ay, S. A., Zimmermann, R., & Kim, S. H. (2008). Viewable Scene Modeling for Geospatial Video Search. Paper presented at the ACM MM08, pp. 309-318.

[5]Ay, S. A., Zimmermann, R., & Kim, S. H. (2010). Relevance Ranking in Georeferenced Video Search. *Multimedia Systems Journal*, 16(2), pp.105-125.

[6]B. Furht.(2008) "Geographic Video Content," *Encyclopedia of Multimedia*, B. Furht, Springer Science+Buisness Media, pp. 271-272.
