Position Paper: A conceptual model of Spatial Video moving objects using Viewpoint data structures.

Paul Lewis[1], Adam Winstanley[1], Stewart Fotheringham[2]

[1]Department of Computer Science/[2]National Centre for Geocomputation,
National University of Ireland Maynooth, County Kildare, Ireland
{paul.lewis, adam.winstanley, stewart.fotheringham}@nuim.ie

Abstract. Spatial Video is any video sequence that has spatial properties associated with it. This paper presents a conceptual model called a Viewpoint which is used to fully define a Spatial Video data set. We highlight how existing models of Spatial Video are limited by their disparate requirements and how our Viewpoint definition can be used in the general case. To do this we define a model that uses the GIS point and polygon data type primitives to create a hypothetically maximum spatial extent for each image contained within a Spatial Video sequence. This concept is shown to both restrict and extend these GIS primitives, but also provide the means to broaden Spatial Video's exposure to a larger number of GIS geospatial analysis operations. Thus, with our conceptual model of Spatial Video as a Viewpoint moving object we can further develop our theories into practical data modelling examples and natural or formal language domains.

1 Introduction

Digital Video recording has become a very prevalent medium in modern society and encompasses many forms, from simple personal camcorders through to sophisticated surveillance systems. In the majority of cases the video footage is usually captured while the device and/or the objects being viewed are in motion. What is of particular interest is when video streams can be, or have been associated with some relevant forms of spatial information such as location, orientation, etc.; to create geographically referenced videographic data. For simplicity, geographically referenced videographic data will be defined as Spatial Video. When such properties are associated with video footage a basic but fundamentally important geographical element has been included, which enables the integration of video with Geographical Information Systems (GIS) [1-3].

However, existing Spatial Video systems are usually bespoke, application specific and not easily modelled into a generalised format that supports a broad range of GIS operations. This situation has resulted in many different methodologies in data collection, processing, storage and use, as well as Terabytes of collected information that is difficult to use outside its original intent. Therefore, we need to find ways to harness these disparate methodologies and large volumes of Spatial Video so they can be integrated within the scope of as many GIS geospatial analysis operations as are relevant. Thus, to provide as general a solution as possible, it is important to use both retrospective and future data sets, along with the many methodologies available to collect them, as a GIS basic data type that satisfies a standardised data structure.

By defining a spatially tagged video stream as a series of Viewpoint data structures we can generalise Spatial Video to describe a number of different fundamental GIS data types that can provide a semantically more meaningful Spatial Video sequence with a broader ability to expand its usage to a greater number of GIS geospatial analysis operations.

2 Spatial Video

In general, video is a well understood concept that can be loosely defined as the set of technological methods and operations used to capture a sequence of moving images. A large number of both hardware and software formats exist for capturing, storing, editing and viewing video. Spatial Video is a specific extension to any of these video format stages that can associate a set of spatial properties to a relevant video stream. These spatial properties can include any number of different parameters that can help define a video frames location, time, orientation etc. Spatial Video's uses are usually limited in nature with many commercial and research groups requiring visual data sets for asset tagging and survey analysis roles [4-6].

Specifically, the Spatial Video data sets being described in this paper contain varying levels of spatial detail; however, the minimum spatial attribute is a GPS track associated to its relevant video stream (Fig. 1. shows an example). The frequencies at which both the video streams and the spatial parameters are captured are dependent upon the hardware. Thus, for a set of GPS tracks that are captured at the standard 1hertz output an appropriate interpolation is usually required to properly relate this data to a video format that may range from 4 to 60 hertz in its recorded frequency.

Fig. 1. This simple example of a partial Spatial Video stream depicts a minimal data set of video and GPS NMEA message relationships.

In essence, the minimal Spatial Video data set is one where each frame of a video stream is indexed to a captured or interpolated GPS NMEA message. While this video stream is self contained and explanatory the GPS NMEA message contains enough spatial information to provide the very basic building block to fully describe the spatial extent of the video. However, this combination of video and spatial properties is normally only used as a GIS enhanced visualisation data source that displays a video sequence based on the location and viewing orientation information that is stored for each captured frame. Thus, the simplest description of a Spatial Video sequence is in this limited role, where the location and orientation of any given video frame can be ascertained. However, it is because of this limited model that we see why Spatial Video has remained restricted in its conceptual understanding as it is not only confined to the afore mentioned role but each SV sequence is self contained and not dynamically searchable or interchangeable. This has led to large volumes of Spatial Video being stored in independent formats and units where each set has no logical or searchable relationship to any other data set.

By redefining our understanding or Spatial Video we believe it is possible to build a more complete conceptual model. To do this a Viewpoint definition of each frame of Spatial Video is proposed, where a minimal set of both spatial and camera parameters can be used to initially understand, then describe and ultimately build a more complete, logical, searchable and usable model for Spatial Video.

3 Viewpoint Concept

To define a Spatial Video model in a broader, generalised GIS context, a move away from the simpler view of knowing the location and orientation of any single video frame is required. By combining the fields of Space Syntax theory and techniques in a GIS [7] with 3D computer graphics, a definition and model can be developed as a Viewpoint spatial data structure. Effectively, a Viewpoint is defined as a refinement of the architectural Isovist or GIS Viewshed concept and the implementation of a viewing Frustum. To do this we build a Spatial Video conceptual description that begins with the most primitive spatial data types; points, lines and polygons, and uses any other spatial parameters to extend or improve this model. These Viewpoints represent each full Spatial Video data set and describe the direct relationship to a single video frame, key frame or set of frames. Ultimately, Viewpoints are a complete understanding of a conceptual model of a Spatial Video data sets object space where its geographical extent is fully described.

To understand these Viewpoints we need to alter our existing Spatial Video model from a frames recorded at a point in space description, to a point in space that has its basic spatial properties both restricted and extended based on the associated image. To explain how this understanding is different we will look at both the restricted and extended nature of a Spatial Video frame point.

3.1 Point Data Type Restriction

When considering the GIS primitive point data type we are describing an entity that has a minimum X and Y parameter in 2D space and a Z parameter if we define a 3D space. Using the basic GPS parameters from a Spatial Video data set we can easily handle the relationship between XYZ and latitude, longitude and altitude. Thus, the point data type can fully describe the spatial location of any image frame, which is the existing model for describing Spatial Video. However, in a general GIS context a point can have many different conceptual meanings when used to represent relevant spatial data sets, but this context can be seen to become fully restricted and definable when incorporated into a Spatial Video data model. So, when we talk about a point representing a frame in a video sequence we can only consider its spatial relevance as an attribute to the video frame because it holds no other spatial information. It can't tell us anything about its spatial environment beyond describing the location of the device that captured the associated image. image/s (Fig. 2 shows an example).

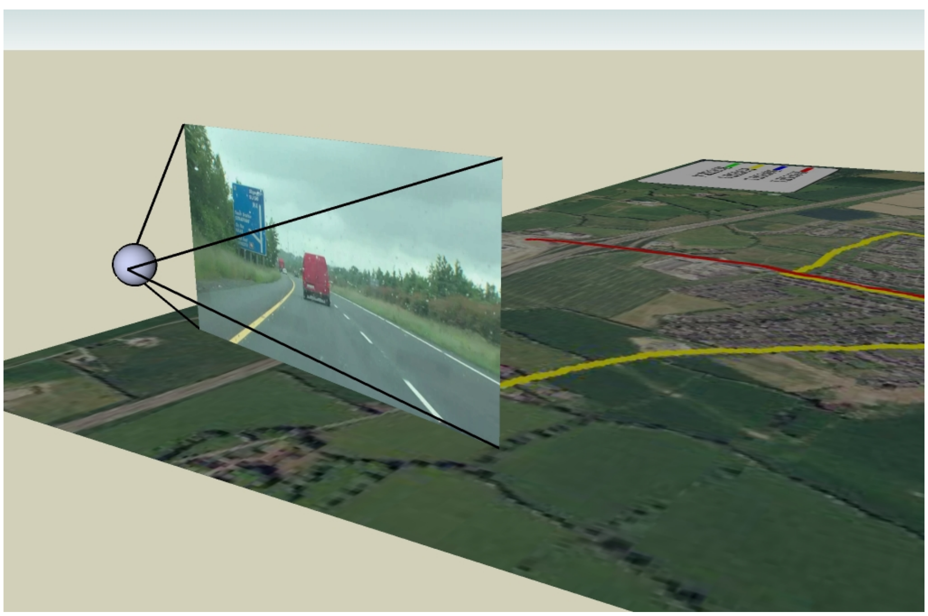

Fig. 2. GIS point data type represented in 3D space. It represents the GPS spatial data and visualizes its relationship to an associated video sequence frame. This representation is a video sequence frame captured from a survey vehicle and displayed in its orthogonal plane to the underlying Ortho Photo. Scales are not preserved.

3.2 Point Data Type Extension

The converse of the point data type restriction is the spatial extent that can be completely defined from the image. An images basic nature is to intrinsically record space and it is this aspect that provides the extension to the point data structure. While it is not necessary to re-define the point data type it is important that the extension be properly handled. Essentially what is being described is a GIS primitive polygon type, or polyhedron in 3D, which fully encloses the geographical extent contained in the images object space. Depending on the camera parameters, this extent is most likely to be shaped as a viewing frustum where the definition of the space will have been offset from the point that defines the image capture location. We therefore are defining a Viewpoint as a GIS primitive point data type that provides an origin, probably offset, for the construction of a polygon data type (Fig. 3 shows and example).

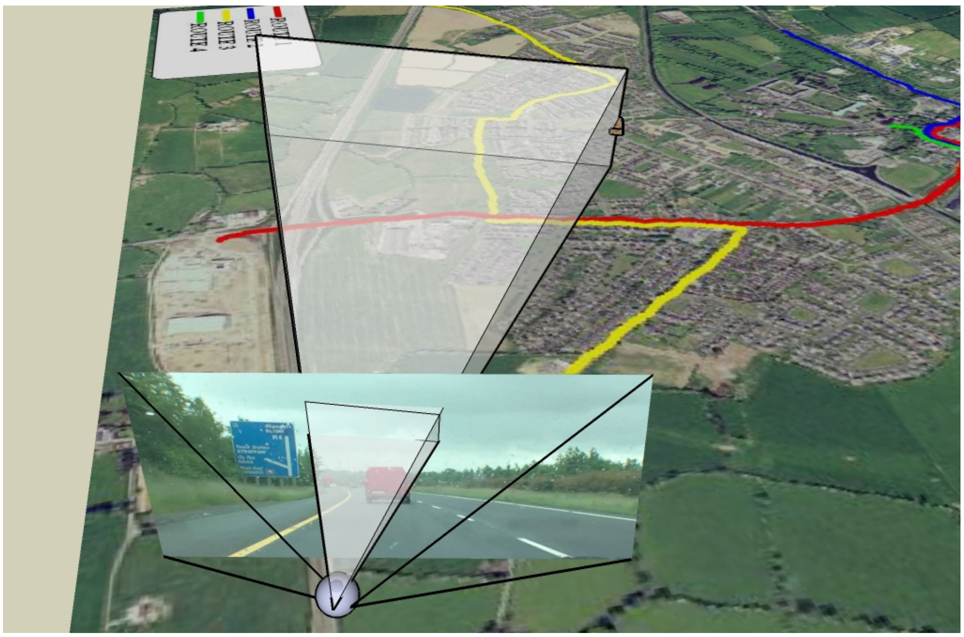

Fig. 3. Visualisation of a Viewpoint in 3D space and how it conceptually relates to a video sequence frame and GPS point. While the image defines a viewing plane that is orthogonal to the Ortho Photo, in spatial terms the polyhedron or more specifically frustum defines the spatial extent. Scales are not preserved.

4 Viewpoint Operations

In providing this Viewpoint model, what we are hoping is that many different types of spatial query can be performed that return logical video sequences. To explain some of these perceived scenarios we will consider some example queries in 2D space, because of ease of visualisation. The previous sections highlighted the GIS primitive data types of points and polygons that are being used to define a Viewpoint. Thus, any queries involving points, lines and polygons that have existing GIS implementations should be possible and easily related back to a Spatial Video sequence. When performing these operations we are trying to integrate Spatial Video with operations on itself or other, possibly non-video, sources of spatial data. As an example the process of spatial data queries with a Spatial Video data store could involve road centre line or regional polygon data sets.

A simple query involving an arbitrary point's data set could ask questions like 'Return all video sequences originating from these points' or 'Return all video sequences that contain these points'. Both these requests have very distinct operations but subtle meanings. The former would involve a simple query between two point's data sets while the latter would require a GIS contains operation that returns the points contained within the Viewpoints polygons.

Other queries could involve many different and more in-depth questions being asked about a Spatial Video data set, examples could include ‘Given a polygon data set how much of its area has been recorded?’ or ‘Given an area return all footage that views it?’ (Fig. 4. shows an example of this). In this example case, normal spatial operations involving a polygon contains query, would return the three red points, two green points and one yellow point contained within the polygon. However, what is proposed here is that the Viewpoint model will provide the correct semantic understanding for formulating more meaningful spatial queries of this type on a Spatial Video data source. Thus, the restriction and extension nature that Viewpoints impose on the GIS point data structure are realised.

While the three red points are contained within the polygon they are excluded from the result set based on the Viewpoint data structure model. From a Spatial Video perspective these red points provide no information about the polygon, other than being contained within it, as the visual properties defined by the Viewpoint all relate to geographical extents outside the polygon. The, obvious, extension of this model is handling those points that may not be contained within the polygon but define a Viewpoint that in some way is contained. The two green points are said to be fully contained as the complete Viewpoint, both point and polygon, are within the query region. While the three yellow points are said to be partially contained as only parts of the Viewpoint are within the polygon query region.

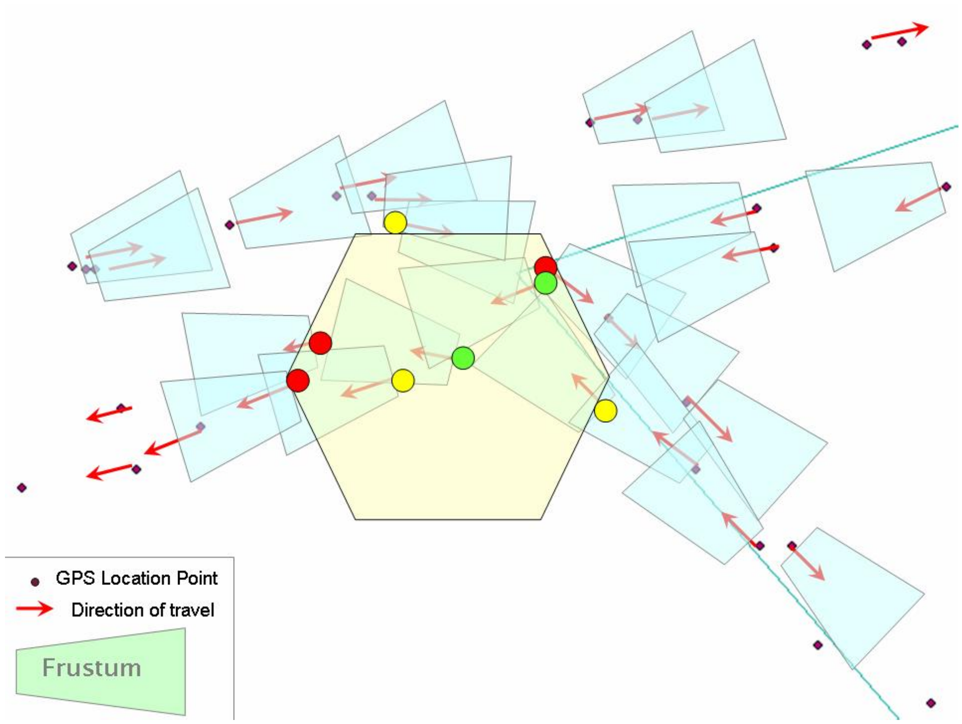

Fig. 4. This diagram displays 5 Spatial Video sequences covering a road junction. Each GPS point, the vehicles direction of travel and a simulated Viewpoint is shown. A sample query involving the yellow polygon should only return valid Video sequences shown by the green and yellow points, the red points, while within the area have no relevance as the spatial extent they represent is not within the query area. Scale is not preserved.

5 Summary

By developing a data structuring system based on a Viewpoints methodology we can facilitate the broader objectives of extending Spatial Videos GIS geospatial analysis functionality. A much simpler and streamlined data store can become central to any indexing, querying and searching system that is dedicated to Spatial Video. This generalised model can easily consider heterogeneous data sets as a single entity and thus provide dynamic cross format access to any relevant Spatial Video sequence. This can be simply done by indexing the Viewpoints against the inherent video format indexing system.

Also, because this data structure is implemented using the point and polygon GIS data type primitives, any GIS geospatial operations can then be preformed that are logical and reasonable when relevant queries involving a Spatial Video sequence are required. Practical implementations of this model are currently developing a spatial database store, with Spatial SQL algorithms being defined that use both point and polygon spatial data types. It is expected that these developments can be used to effectively describe, query and retrieve any Spatial Video data set as a sequence of logical Viewpoints.

References

1. Hirose M., Watanabe S., Endo T.: Generation of wide-range virtual spaces using photographic images. Virtual Reality Annual International Symposium, 1998. Proceedings IEEE 1998. IEEE, Atlanta, GA, USA (1998) 234-241
2. Kawasaki H., Yatabe T., Ikeuchi K., Sakauchi M.: Automatic modeling of a 3D city map from real-world video. International Multimedia Conference, Vol. Proceedings of the seventh ACM international conference on Multimedia (Part 1). ACM Press, Orlando, Florida, United States (1999) 11-18
3. Kim Kyong-Ho, Kim Sung-Soo, Lee Sung-Ho, Park Jong-Hyun, Lee Jong-Hun: The interactive geographic video. Geoscience and Remote Sensing Symposium, 2003. IGARSS '03. Proceedings. 2003 IEEE International, Vol. 1, Toulouse, France (2003) 59-61
4. Spatial video solutions for transportation sector., www.routemapper.net
5. Spatial Video hardware and software website, http://www.redhensystems.com
6. Business shop front spatial image location and direction system, http://maps.a9.com/?
7. Jiang B., C., C.: Integration of space syntax into GIS: new perspectives for urban morphology. Transactions in GIS **6** (2002) 295-309