Korean Journal of Remote Sensing, Vol.19, No.1, 2003, pp.71~80

Web-Based Media GIS Architecture Using the Virtual World Mapping Technique

Sung-Soo Kim, Kyong-Ho Kim, and Kyoung-Ok Kim

Spatial Information Technology Center, Electronics and Telecommunications Research Institute

Abstract : In this paper, we propose web-based Media GIS architecture using 3D geographical database and GPS-related data resulted from 4S-Van. We introduce a novel interoperable geographical data service concept; so-called, Virtual World Mapping (VWM) that can map 3D graphic world with real-world video. Our proposed method can easily retrieve geographical information and attributes to reconstruct 3D virtual space according to certain frame in video sequences.

Our proposed system architecture also has an advantage that can provide geographical information service with video stream without any image processing procedures. In addition to, describing the details of our components, we present a Media GIS web service system by using GeoVideoServer, which performs VWM technique.

Key Words : Video GIS, 3D GIS, Media Technology, Data Interoperability.

1. Introduction

An increasing number of users of GIS have been asking for 3D and multimedia (video, audio etc.) extensions for years. Surprisingly, few convincing systems have been implemented yet. The problems can be summarized in three points: building 3D geo-spatial database, constructing real-world multimedia data, especially video and audio streams, and providing an interface to visualize and interoperate each other efficiently.

Our research aim is to define a system connecting geo-referenced video material with the geographical information of the real-world space; this information may be stored in one or different sources. In our research case, we consider that video material and 3D geographical information is stored independently to improve data interoperability. User can navigate video stream and 3D virtual space by geographical contents (e.g., GPS and camera data) and retrieve attributes of certain geographic facilities. This is what we have called the Media GIS system.

As geographical information is still strongly 2D oriented, which is mostly due to data capture capabilities, a 3D GIS must allow to import, process, visualize and export 2D data. A traditional GIS provides only 2D representation of the spatial entities using simple primitives of points, lines and polygons.

2.5 D data consists of 2D topology and 3D feature

Received 3 January 2003; Accepted 10 February 2003.

geometry. A 2.5D GIS add a third dimension to the display by adding a third dimension using one of the attribute information such as elevation, land use, land cover, rain fall, etc. The attribute property of the object remains constant along the third dimension.

The recent research trend is to build a truly 3D GIS (Virtual GIS). Current research developments include navigation in a virtual environment and simple querying of the attribute database.

Recently, research activities of Video GIS, which provides geographic information or makes decision by using real-world video streams, are progressing. The Media GIS, which discuss in this paper, is a GIS technology that provides a geographic information services by mutually connecting two-dimensional map and 3D database with real-time video streaming.

The key issue in current Media GIS implementation is how to interoperate and link among various heterogeneous data sources.

By solution for data interoperability, we can implement a GIS system by using OLE/DB technology in Microsoft. But the solution for data linkage of heterogeneous data sources with uni-direction or bi-direction is not presented yet except method through whole data integration.

Until now, Data, which have been used as typical data format in GIS systems, are two-dimensional vector map, three-dimensional database and GPS-related data acquired from 4S-Van. To interoperate these heterogeneous data, 4S technology is recently introduced.

Let's examine bi-directional geographic information service concept among various spatial datasets. For example, let's suppose that GIS users want following two geographic information services. In one case, GIS users want to retrieve geographic information of arbitrary facility according to certain frame in video sequences with navigation. In the other case, GIS users want to browse the video clip according to the user requested query from 2-dimensional vector map. If any system satisfies above two cases, then we can say that these systems provide bi-directional geographic information service. But, most GIS systems are difficult to provide data interoperability and bi-directional geographic information services due to their different geographic information data format.

The ultimate goal of this research is providing geographic information services, which interoperate heterogeneous data that have different spatial data formats and attributes without any special data correction and integration works.

We introduce a novel interoperable geographical data service concept; so-called, Virtual World Mapping (VWM) that can map 3D graphic world with real-world video. The details of this algorithm will be presented in Section 3.

We here summarize some of its features:

Interoperability By designing and embodying a standardized data provider, we can solve the problem resulted from different aspect of data format and database that used in current commercial GIS product.

Standardization We consider in standardization of three-dimensional geometric model as extending the two-dimensional geometric model proposed by OpenGIS Consortium (OGC).

Reusability We can improve the software reusability by constructing several core components such as data provider, data manager, scene manager component.

Adaptability We design the scene manager component to easily adapt in 3D GIS applications as well as various application fields such as CAD/CAM, computer graphics and virtual reality applications.

Performance We also use Well-Known Binary (WKB) to store geometry information as in OGC simple features specification (OGC, 1999).

2. Related Works

Most of the current geographic information systems implemented so far are concerned with 2D or 3D geographic visualization and analysis. But, few convincing systems for video geographic information services have been implemented yet.

We introduce a new approach for video geographic information services by interoperating GPS-related data with 3D GIS database.

The Aspen Movie Map Project, developed at MIT in 1978, is historically the first project combining video and geographical information (Lippman *et al.*, 1980). Using four cameras on a truck, the streets of Aspen were filmed (in both directions), taking an image every three meters. The system used two screens, a vertical one for the video and a horizontal one that showed the street map of Aspen. The user could point to a spot on the map and jump directly to it instead of finding the way through the city.

Many projects have used video clips in a similar way. The most typical case is multimedia atlases where the user can find video clips of locations or providing a deeper definition of any geographical concept.

Other applications with a geographical background have used video clips: a collaborative hypermedia tool for urban planning. Most of these systems simply link 2D vector maps with video clips.

Recently, Navarrete (Navarrete *et al.*, 2001) proposed a method, which performs image segmentation about a certain video frame through image processing procedures for combining video and geographic information.

The main problem of this method when dealing with big sources of video is how to segment it, i.e. how to choose the fragments of video that will be the base of later indexing and search. On option is a handmade segmentation of video, but this is too expensive for huge archives.

Moreover, manual indexing has other problems as Smeaton (Smeaton *et al.*, 2000) points to :

- No consistency of interpretation by a single person over time
- No consistency of interpretation among a population of interpreters
- No universally agreed format of the representation, whether keyword, captions or some knowledge-based information.

Due to these reasons, automatic segmentation of video has been an intensive research field in the late years.

In 3D GIS, we had proposed scene modeler component, which can model buildings or roads by using 2D geometry information with 3D attributes such as a height of building, a width of road, to improve the data reusability and interoperability.

We classify three possible approaches for constructing a video GIS system by using various spatial information processing technologies.

The first approach is based on 4S-Van technique, which performs stereo image processing for two CCD images (left and right image) result from the 4S-Van to construct three-dimensional spatial information (Lee *et al.*, 2001; Joo *et al.*, 2001).

4S represents four systems that are commonly related to spatial information: Geographic Information System (GIS), Global Navigation Satellite System (GNSS), Spatial Imagery Information System (SIIS), and Intelligent Transport System (ITS).

4S-Van technique is a system for spatial data construction that is heart of 4S technology. It enables real-time acquisition of the position information and accurate image data in remote site.

The disadvantage of this approach is that it requires manual image processing of these two images to construct three-dimensional geometry and attributes. And it is difficult to provide data interoperability among 2D, 3D and GPS-related data.

The second approach is based on MPEG-4 standard encoding, which encodes spatial objects in every video frame according to MPEG-4 scene representation format. This approach also requires a lot of manual processing for the MPEG-4 authoring.

The third approach is based on 3D GIS database with GPS-related data, which will propose in this paper. We propose Media GIS system architecture, which provides geographical information of geo-features in video sequences by using 3D geographical database based on 2D geometry and GPS-related data resulted from 4S-Van. Our proposed system architecture also has an advantage that can provide geographical information service with video stream without any image processing procedures unlike other approaches.

The major contribution of this work is to present a new paradigm of video geographic information service by providing data interoperability among heterogeneous spatial data.

3. Media GIS System

1) System Architecture

We call our proposed system as Media GIS to meaning that can interoperate various visualization methods of spatial data such as 2D, 3D and multimedia data (Fig. 1).

The Media GIS system consists of GeoVideoServer for geographic information linking and GeoVideoClient for video browsing and attribute query. Fig. 2 shows the component diagram of proposed Media GIS system.

The GeoVideoServer consists of the ETRI's Distributed Spatial Engine (EDGE) components such as data provider component, data manager component and scene manager component. The data provider component provides data access functionalities through common interfaces from different data sources.

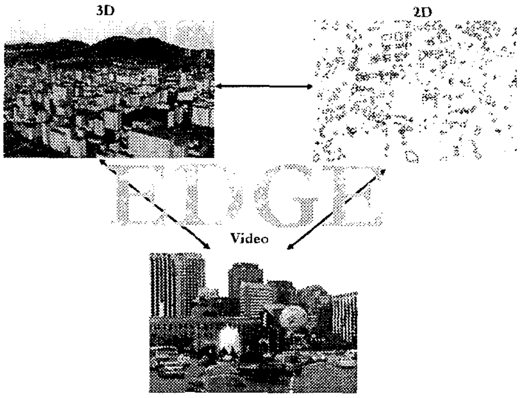


Fig. 1. Unified visualization framework of the EDGE (ETRI's Distributed Geospatial Engine).

In simple features specification of the Open GIS Consortium (OGC), the Well-Known Binary (WKB) representation for geometry provides a portable representation of a geometry value as a contiguous stream of bytes (OGC 1999). We also use WKB representation to store geometry information as in OGC simple features specification (Lee *et al.*, 2001).

The data manager component manages data that received from data provider as an internal form according to system architecture. The scene manager component creates and manages the scene graph for 3D rendering and visualization.

We also define remote procedure call (RPC) interface to communicate GeoVideoServer and GeoVideoClient

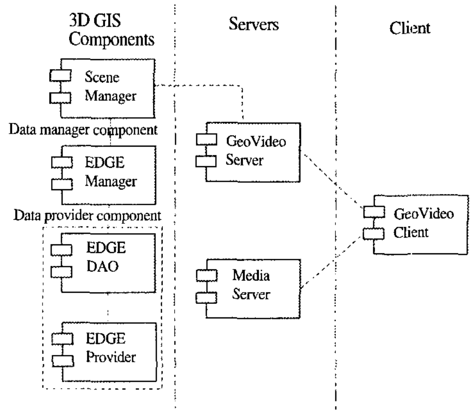


Fig. 2. Media GIS system architecture.

```
void RequestFeatureInfo(
[in] unsigned char RPC FAR *pszVideoFilename,
[in] PointInt ptPickPoint,
[in] long lFrameNo,
[in] long lVideoWidith,
[in] long lVideoHeight,
[out] long RPC FAR *lSelectedID );
```

Fig. 3. RPC interface for client/server communication.

each other (Fig. 3).

There are five input parameters for client and server communication such as a video filename, picking position, a current video frame number, a width and a height of video. The GeoVideoServer passes selected feature ID, which are obtained from picking operation, to client by using RequestFeatureInfo method according to client inputs.

2) Spatial Data Construction

There are two kinds of spatial data that is used in Media GIS system. One is a 3D GIS database that consists of 2D vector geometries and additional 3D attributes. The other is video stream with GPS-related data, such as camera position, camera internal and external parameters, which is obtained from 4S-Van.

Hardware architecture of 4S-Van consists of data store part and sensor part. Sensor part has global positioning system (GPS), inertial measurement unit (IMU), color CCD camera, B/W CCD camera, and infrared rays camera. We also use a video camera to construct video streams with GPS-related data through revision between video and CCD image. We perform spatial data modeling which has video and GPS camera data as following.

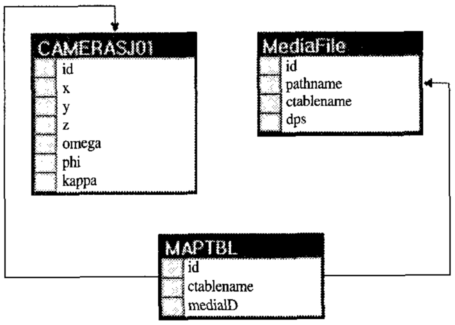


Fig. 4. Media database diagram.

Fig. 4 shows table relationships in our Media database for video geographic information services. Media database has camera tables, media file table and mapping table. The camera table includes camera position P_C (x_c, y_c, z_c) and orientation O_C (ω_c, ϕ_c, κ_c).

We use geometry information of 2D vector map and additional 3D attributes such as base DEM height of building, height of building to construct 3D GIS database.

3) Virtual World Mapping

We introduce a new method, so-called VWM, for linking geo-feature in video and geographic information in the Media GIS. The Virtual World Mapping (VWM) method uses ray-box intersection algorithm for virtual city constructed from 2D geometry and additional 3D attributes.

The data manager component has the LOD (Level-of-Detail) modeler library which performs 3D geometric modeling by using synthetic modeling concept (Kim *et al.*, ISRS 2001). Our LOD modeler library coarsens or fines the 3D geo-features (i.e. building, road) according to the request of user or client program.

There are two important system elements to model the static LOD geo-feature model efficiently. One is the rule-based modeling engine and the other is model library. The rule-based modeling engine refers a rule parameter description file which is saved as a XML file form.

For example, in case of a building, our proposed modeling system can create 3D polygonal model to process 2D profile geometry of a building and height attributes as an input by using the rule-based modeling engine and

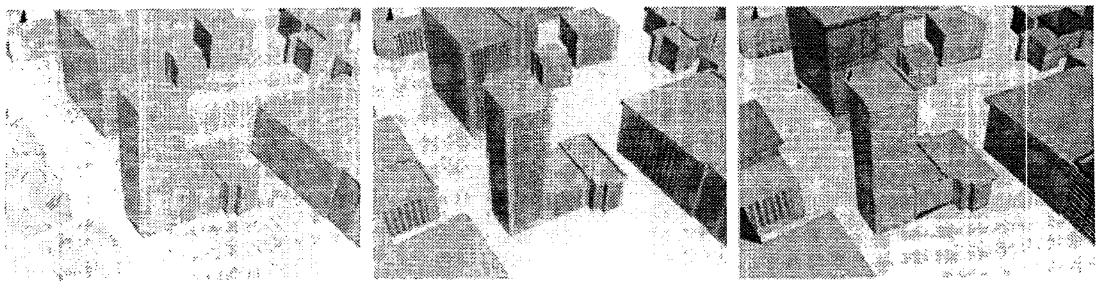

(a) (b) (c)

Fig. 5. LOD modeling results of the synthetic modeling. ((a) τ(*LOD level*)=1, (b) τ=2, (d) τ=3).

model library. Fig. 5 shows LOD modeling results of synthetic modeling according to the LOD level (τ).

The GeoVideoServer waits for user input to perform VWM after rendered 3D building models. Fig. 7 shows the processing flow of VWM technique for user query of the GeoVideoClient.

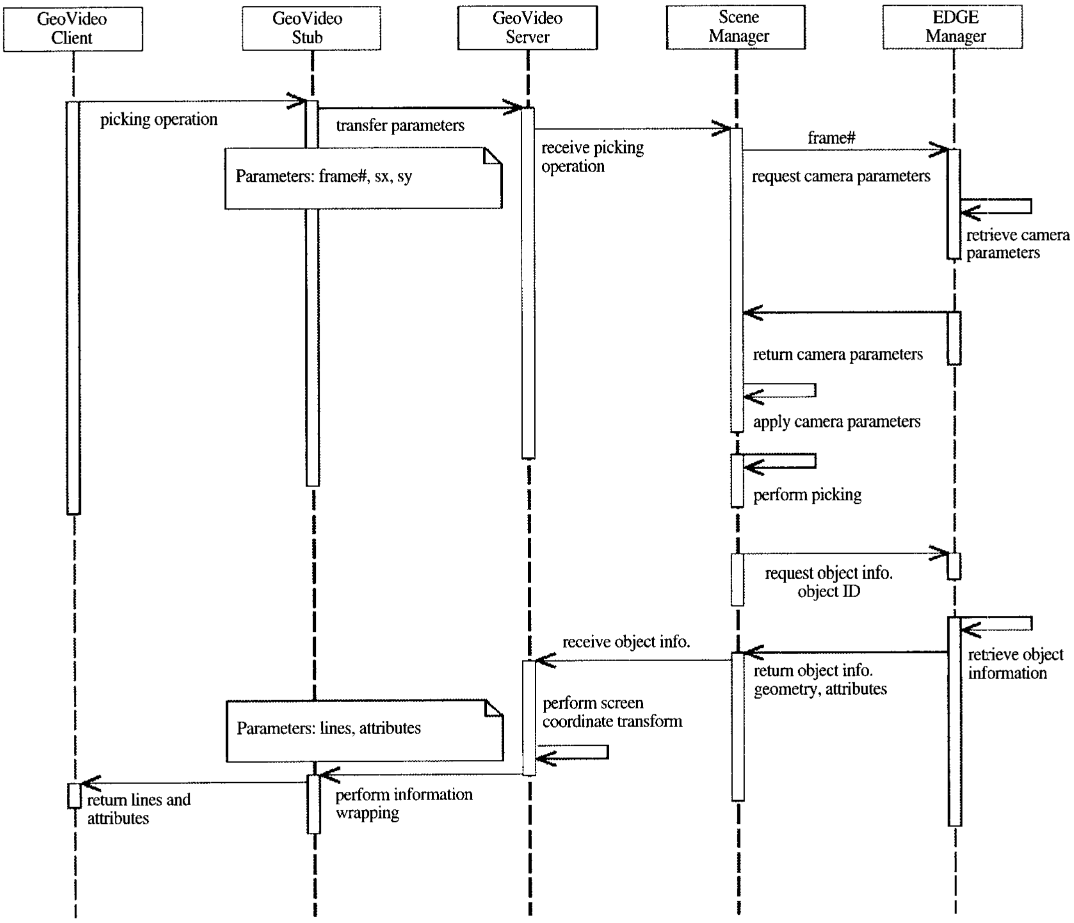


Fig. 6. Sequence diagram of the VWM procedure.

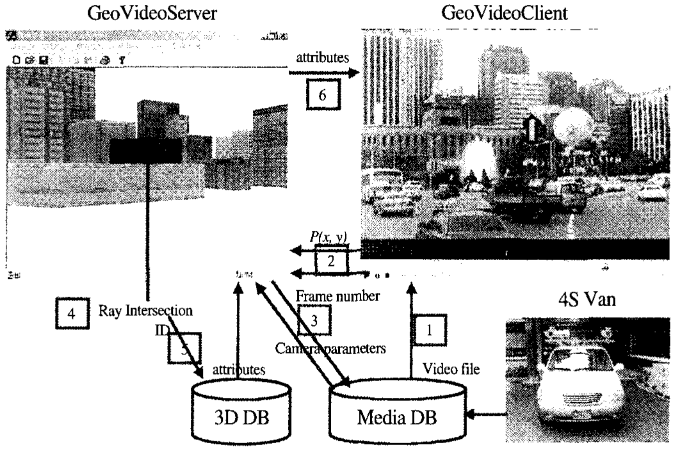


Fig. 7. \VWM processing flow.

The query processing flow for GeoVideoClient is detailed below.

1. First, the GeoVideoClient passes the current frame number Fn in selected video sequence and mouse-clicked position $P(x, y)$ to GeoVideoServer.
2. The GeoVideoServer gets a camera data; position P_c (x_c, y_c, z_c) and orientation O_c $(\omega_c, \phi_c, \kappa_c)$ from the media database according to the video filename and current frame number.
3. The GeoVideoServer locates camera position to P_c (x_c, y_c, z_c) with O_c $(\omega_c, \phi_c, \kappa_c)$ in 3D virtual space and then performs picking operation at the $P(x, y)$ to get the identification (ID) of relevant facilities through computing ray and bounding box of facility intersection.
4. Finally, the GeoVideoServer passes the attributes of selected facilities in the 3D database to the GeoVideoClient according to the selected ID.

The VWM technique computes ray-box intersection to perform the picking operation.

To check if a ray intersects such a box is straightforward. We treat each pair of parallel planes in turn, calculating the distance along the ray to the first plane (t_{near}) and the distance to the second plane (t_{far}).

The larger value of t_{near} and the smaller value of t_{far} is retained between comparisons. If the larger value of tnear is greater than the smaller value of t_{far}, the ray cannot intersect the box (Alan *et al.*, 2001).

This is shown, for an example, in the xy plane in Fig. 8. The Ray A is not intersected otherwise Ray B is intersected. If a hit occurs then the intersection is given by t_{near}.

$t_{near(larger)} > t_{far(smaller)}$: not intersected
$t_{near(larger)} < t_{far(smaller)}$: intersected

Our implemented picking operation can require Θ(n) time because of it must perform operations for all bounding box in the 3D scene, where n denotes the total number of buildings.

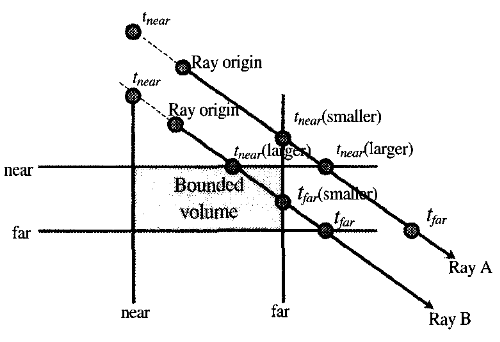


Fig. 8. Ray-box intersection.

Algorithm 1 contains the pseudo-code for the GeoVideoServer-side procedure.

Algorithm 1. RequestFeatureInfo

```
// get GeoVideoServer view
CGeoVideoServerView* pView = GetServerView();
// resize view according to the video size.
pView→resizeView(lFrameWidth, lFrameHeight);
// set the camera position of the requested video frame.
pView→setCamera(lFrameNo);
// compute ray-box intersection.
pView→pick3DSpace(ptPickPoint.x, ptPickPoint.y);
// get attributes of selected feature.
pView→getLinesAttributes(lFrameNo, lLineCount,
ptDeviceLine, lDescStringLen, Desc);
```

The pick3DSpace member function computes ray-box intersection. For Media GIS web services, we use the Microsoft Media Server for media file streaming service. We implement Media GIS web service page using the GeoVideoClient as an ActiveX control and ASP (Active Server Page). Fig. 9 illustrates the proposed Media GIS web service architecture.

4. Experimental Results

We have implemented our VWM algorithm in C++ and OpenGL library, tested our implementation on several datasets which are obtained from Jung-Gu, Seoul, Korea by using the 4S-Van.

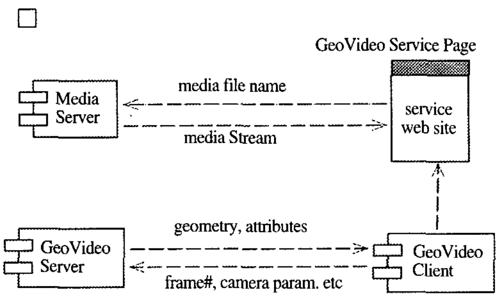


Fig. 9.Media GIS web service architecture.

All experiments were conducted on a 2.4GHz Pentium IV running Microsoft Windows 2000 Server and Professional. Our GeoVideoServer and Client systems had 1GB RAM and a graphics accelerator based on the NVIDIA GeForce 256 chip. We also use the Internet Information Server (IIS) and Microsoft Media Server (MMS) for the GeoVideoServer. The GeoVideoClient is implemented in C++ and ATL/COM as an ActiveX control.

Traditional RPC communication has the firewall problem. We overcome this problem by using ISAPI (Internet Server API) Extension for GeoVideo client/server communication.

The error range of GPS data is different each other according to GPS devices. Generally, the error of the altitude is greater than that of latitude and gradient. The altitude error is about 5-20m (Byun *et al.*, 2001).

In our work, we correct the altitude error to 15m for our experiments. We also correct GPS-related data by applying internal parameters (i.e. focal length, FOV; field of view) and external parameters (i. e. P_c (x_c, y_c, z_c), O_c (ω_c, ϕ_c, κ_c)) of a video camera.

We perform picking operation at least 70 random points in the GeoVideoClient (Fig. 10) to measure the overall query processing accuracy. Then we can get a matrix, called an error matrix as the Table 1.

The overall picking accuracy can be computed as the total number of correct solutions (the sum of the diagonal

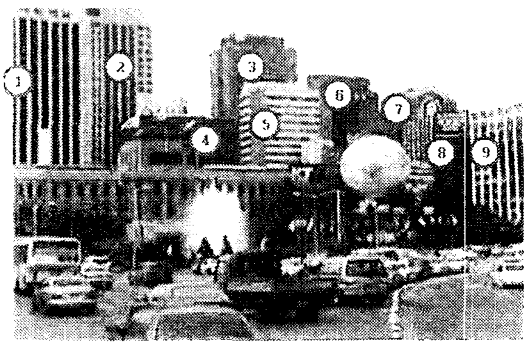


Fig. 10. Test region for picking accuracy measure.

Table 1. Error matrix for picking accuracy.

in/out	1	2	3	4	5	row total
1	**80**	24	0	0	0	104
2	4	**60**	0	0	0	64
3	0	0	**50**	0	2	52
4	0	0	0	**62**	4	66
5	0	0	24	16	**64**	104
column total	84	84	74	78	70	**390**

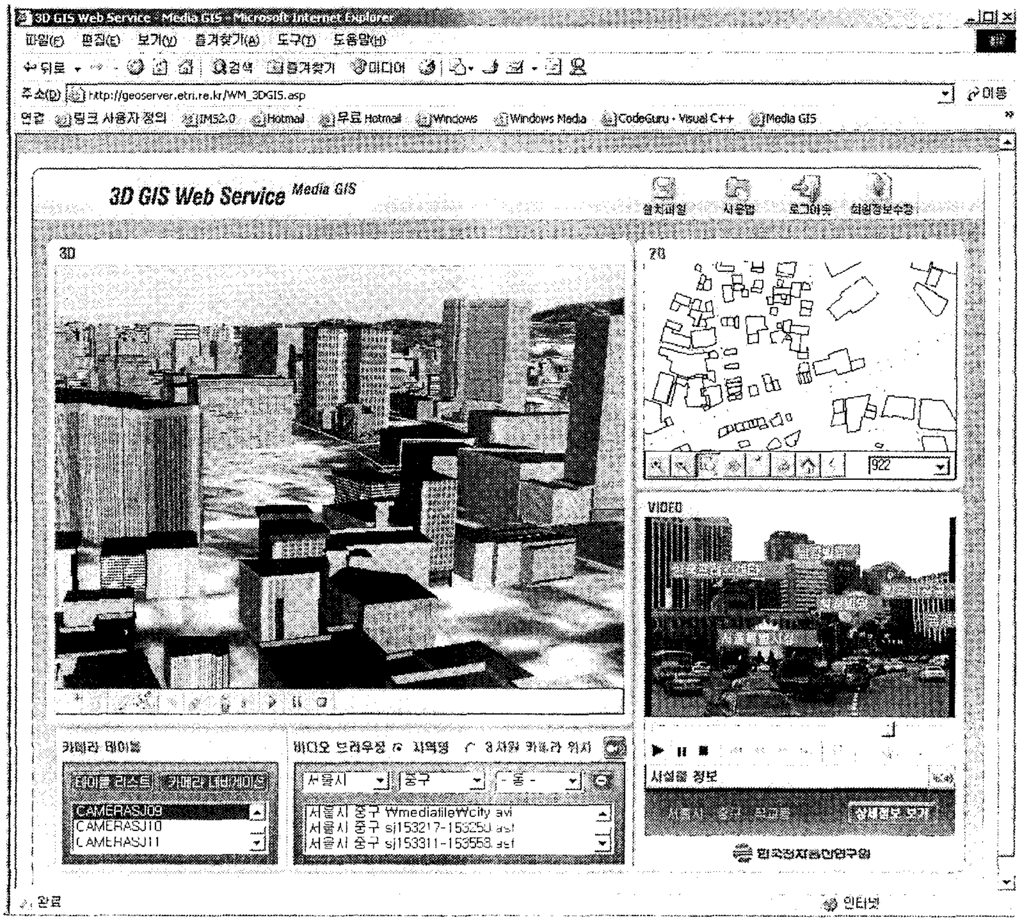
cells) divided by the total number of cells. The overall picking accuracy is (80+60+50+62+64)/390, or 81 %.

The accuracy decreases at the building number 3 and 4 because of we have rendered 3D buildings using only 2D geometries and 3D additional attributes such as height of building. And it is difficult to augment video with geometries of 3D virtual space due to the data acquisition error.

Fig. 11 shows the Media GIS system running on the internet environment.

5. Conclusions

We have presented a new web-based Media GIS web architecture using 3D geographical database and GPS-related data resulted from 4S-Van. We introduce a novel interoperable geographical data service concept; so-



http://geoserver.etri.re.kr/WM_3DGIS.asp

CAMERASJ09
CAMERASJ10
CAMERASJ11

서울시 중구 WmediafileWcity avi
서울시 중구 sj153217-153250.asf
서울시 중구 sj153311-153553.asf

Fig. 11. Media GIS web GUI (graphic user interface).

called, Virtual World Mapping (VWM) that can map 3D graphic world with real-world video.

Our proposed method can easily retrieve geographical information and attributes to reconstruct 3D virtual space according to certain frame in video sequences.

The major contribution of this work is to present a new paradigm of video geographic information service by providing data interoperability among heterogeneous spatial data.

Our proposed system architecture also has an advantage that can provide geographical information service with video stream without any image processing procedures.

Future works include:

- Automatic algorithm for error correction for camera parameters.
- Augmented reality technology for mixing 3D virtual world with real-world video.

Acknowledgements

This study was conducted as a part of MIC project "3D GIS software".

References

Watt, A. and F. Policarpo, 2001. *3D Games: Real-time Rendering and Software Technology*, Addison-Wesley, pp. 14-22.

Joo, I. H., M. S. Kim, and B. W. Oh, 2001. Development of Disaster Control System Based on 4-S VAN and Mobile Environment, *Proc. of International Symposium on Remote Sensing*, pp. 137-142.

Byun, K. J., D. C. Lee, J. W. Kim, C. Y. Park, and B. T. Jang, 2001. A Study on Car Navigation System Based on Augmented Reality Technology, *Proc. of Korean Information Science Society Fall Conference*, pp. 493-495.

Lippman, A., 1980. Movie Maps: An Application of the Optical Videodisc to Computer Graphics, *Proc. of SIGGRAPH'*80, pp. 32-43.

Open GIS Consortium, 1990. OpenGIS Simple Feature Specification for OLE/COM, OpenGIS Implementation Specifications, Revision 1.1.

Lee, S. H., K. H. Kim, S. K. Choe, S. S. Kim, and Y. K. Yang, 2001. Extension of OpenGIS OLE/COM SFS for 3-Dimensional GIS, *Proc. of International Symposium on Remote Sensing*, pp. 248-251.

Lee, S. Y., B. W. Oh, and E. Y. Han, A Study on Application for 4S-Van, 2001. *Proc. of International Symposium on Remote Sensing*, pp.124-127.

Smeaton, A.F, 2000. Browsing and Searching Digital Video and Digital Audio Information, *Proc. of the 3rd European Summer School on Information Retrieval* (*ESSIR*2000).

Kim, S. S., S. K. Choe, J. H. Lee, and Y. K. Yang, 2001. Rule-Based Modeler Component Design for 3D GIS Software, *Proc. of International Symposium on Remote Sensing*, pp. 89-94.

Kim, S. S, S. K. Choe, J. H. Lee, and Y. K. Yang, 2001. The Design of Scene Modeler Component for 3D GIS Software, *Proc. of Korean Information Processing Society Fall Conference*, pp. 81-84.

Kim, S. S., K. S. Kim, S. H. Lee, S. K. Choe, K. H. Kim, J. H. Lee, and Y. K. Yang, 2001. The Three-Dimensional Extension for Mapbase Components using the Rule-Based Modeling, *Proc. of Korean Multimedia Society Fall Conference*, pp. 171-176.

Navarrete, T., J. Blat, 2001. VideoGIS: Combining Video and Geographical Information, Research Report, Dept. of Computer Science, University of Pompeu Fabra.