GEOVISUALIZATION FOR SMART VIDEO SURVEILLANCE

Reinier Oves García, Luis Valentín, Sergio A. Serrano, Miguel A. Palacios-Alonso, L. Enrique Sucar

Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE)
Sta. María Tonantzintla, CP 72840, Puebla, Mexico.
{ovesreinier, luisnvnc, sserrano, mpalacio, esucar}@inaoep.mx

KEY WORDS: Smart surveillance, Smart sensing, Smart security, Geovisualization, FIWARE

ABSTRACT:

Nowadays with the emergence of smart cities and the creation of new sensors capable to connect to the network, it is not only possible to monitor the entire infrastructure of a city, including roads, bridges, rail/subways, airports, communications, water, power, but also to optimize its resources, plan its preventive maintenance and monitor security aspects while maximizing services for its citizens. In particular, the security aspect is one of the most important issues due to the need to ensure the safety of people. However, if we want to have a good security system, it is necessary to take into account the way that we are going to present the information. In order to show the amount of information generated by sensing devices in real time in an understandable way, several visualization techniques are proposed for both local (involves sensing devices in a separated way) and global visualization (involves sensing devices as a whole). Taking into consideration that the information is produced and transmitted from a geographic location, the integration of a Geographic Information System to manage and visualize the behavior of data becomes very relevant. With the purpose of facilitating the decision-making process in a security system, we have integrated the visualization techniques and the Geographic Information System to produce a smart security system, based on a cloud computing architecture, to show relevant information about a set of monitored areas with video cameras.

1. INTRODUCTION

Ensuring society’s security and welfare is one of the main concerns of every government, however, this is not an easy task. The main purpose of video surveillance is to monitor activity occurring within an area of interest. So far, video surveillance systems have proved to be very good for access control, preventing crime, or even when criminal activity has occurred, video surveillance turns to be very helpful on providing valuable evidence for further criminal investigations.

Although video surveillance systems have been widely used for several years, there are some issues closely related to this type of application. Problems start to arise when the amount of video cameras connected to the system becomes considerably large, this decreases the system’s overall monitoring performance. The greatest issue on assigning a security guard for the task of watching over a set of video streams, is that people’s attention, in general, is limited; therefore, there is always a good chance that they miss some relevant information, that in the context of security, is considered inadmissible because of the consequences this might have. Another important problem with classic surveillance systems is that the data stored by the system (video footage) has such enormous dimensions, that it becomes intractable for off-line analysis; for instance, looking for evidence on some reported criminal activity. Even when the problem of intractability in analysis of information for video surveillance systems has been well defined, approaches on solving this issue are still in an early stage.

On the other hand, Scientific Visualization (SciVis) aims to transform data, through coding techniques, in a perceptible representation that improves the understanding and communication of models, at the same time that increases the effectiveness in its analysis and evaluation. Visualization techniques, in general, are used for exploratory analysis, confirmatory analysis and presentations. In the last two decades there has been an increase in the use of information systems and scientific visualization that has facilitated the analysis of large volumes of data (Bolstad, 2005, Kronenfeld, 2006). Also, another technology that has been applied in different research areas along with Scientific Visualization are the Geographical Information Systems (GIS), which are based on the integration of high performance computers, geospatial databases and visualization software for editing, managing and deploying geographical data in different ways.

As previously stated, the main problem in video surveillance systems is the intractability in analysis, caused by the large amounts of data generated by the system and people’s limited capability to perform several tasks simultaneously. Nevertheless, there has been a lot of work made on the development of computer vision techniques to detect events from a real-time video stream. The automatic detection of certain situations, by its own, does not solve the intractability issue. A method able to represent huge amounts of information into compact and understandable graphics is still required. For this reason, Scientific Visualization becomes a very useful tool to address the intractability problem by generating images that show, with great quality, as many aspects of the original information as possible. Moreover, by integrating a Geographical Information System along with Scientific Visualization to a video surveillance system, besides having a synthesized version of the data, it helps the user to know where the data is coming from, by relating every data plot (Scientific
Visualization) to a geographical location.

In this work we propose a video surveillance system that integrates GIS technologies and Scientific Visualization techniques. With this system we are able to acquire a video stream from a set of cameras, process that information and display it in such way that we give to the user access to different layers of information in order to provide a means that makes easy to analyze and understand what is happening in a monitored area.

The rest of the paper is organized as follows. Section 2 summarizes the main advances in geovisualization, smart cities and smart surveillance. Section 3 describes the proposed surveillance system aided by geovisualization. In Section 4 is presented an application of the proposed system, and finally in Section 5 is presented the conclusions and future work.

2. RELATED WORK

In this section we summarize the previous work on geovisualization and visualization for video surveillance.

2.1 Geovisualization

Geovisualization is a multidisciplinary field which involves Mapping, SciVis, Scientific Analysis, Image Analysis, Information Visualization, Exploratory Analysis and GIS aiming to provide methods and tools for visual exploration and analysis of Georeferenced data (Vázquez R., 2016).

There are two common tendencies in GIS development: Desktop-based and Web-based. Desktop-based GIS are characterized by its high computing power and visualization capabilities. ArcGIS (Johnston et al., 2001), GRASS GIS (GIS, 2011), QUANTUM GIS (Hugentobler, 2008), Geovista Studio (Takatsuka and Gahegan, 2002) and gvSIG (Anguix, 2009) are examples of Desktop-based GIS. On the other hand, a Web-based GIS is mainly focused in the collaborative analysis of geographical information through map servers. For instance, Mappoint (Kropa, 2006) and Geoerver (Iacovella, 2013) are web-based map servers, while Google maps (Miller, 2006) and OSM maps (Haklay and Weber, 2008) are considered GIS.

Advances in integrating GIS and SciVis techniques result in the creation of systems like Snap-Together Visualization (North and Shneiderman, 2000), Geovista Studio, VIS-Stamp (Guo et al., 2006) and GAV Flash Tools (Jern et al., 2010) among others. All these tools enable the visualization of temporal data, gathered by experts or sensing devices (e.g. temperature, rain, etc.), in the exact location where they were measured.

Different visualization techniques have been used to improve data understanding of a specific phenomena. The most common, easy and understandable way to visualize data is through classical charts (pie-charts, bar-charts, time-series, etc.) (Andrienko et al., 2003) and heatmaps (Pryke et al., 2007). On the other hand, more complex information, such as multiparametric data, requires more complex techniques (multiparametric data visualization techniques) (Keim, 2000, Inselberg and Dimsdale, 1990, Cleveland, 1993, Keim, 2001). Circle segments, Recursive patterns, Parallel coordinates and Chernoff face are examples of these.

During the last two decades, Geovisualization has been used for different applications such as exploring traffic data (air, ground), transportation routes, diseases, weather, etc. Thus, by combining Geovisualization with video surveillance gives us some advantages. For instance, we have the possibility to develop, implement and embed ICTs (Information and Communication Technologies) into a Smart City and consequently provide the user access to different layers of information that might be combined and integrated to facilitate the analysis of a situation, and improve the decision making performance (Daniel and Doran, 2013).

2.2 Visualization for video surveillance

In the past couple of decades, video surveillance applications have generated a lot of interest, due to the wide variety of problems they have been applied successfully on, e.g., access control, preventing crime and monitoring public places in general. However, because in most of the cases a human operator is assigned to watch over a set of video streams, and considering the large dimensions of data generated by a video surveillance system, there is a problem of intractability. For this reason, several works have been done to solve this issue. For instance, in (Daniel and Chen, 2003) a method for summarizing video sequences is proposed, using volume visualization techniques they demonstrate the effectiveness of their method to convey meaningful information contained in video sequences.

In (Sebe et al., 2003) they developed a visualization method based on augmented virtual environments that fuses dynamic imagery with 3D models to help observers to comprehend multiple video streams. In (Irani et al., 1995) they describe the use of mosaic images to represent the information in video sequences. (DeCamp et al., 2010) developed a visualization system, which they call HouseFly, that synthesizes audio-visual recordings from multiple sensors, and also meta-data generated from the original information. Another interesting video visualization technique that differs from the ones seen so far is proposed by (Goldman et al., 2006), where they build a storyboard from multiple input frames which are annotated using outlines, arrows and descriptive text. The advantage their method has over standard video representations, is that it requires just a moment to observe and comprehend, while it retains a significant amount of detail from the video.

However, when it comes to video surveillance systems, it is almost impossible for a user to keep track on what is occurring within a large set of video streams simultaneously. In the following section, we propose a video surveillance system that integrates Geovisualization techniques to provide a representation of data in a way that the exploration, analysis and understanding of all of the monitored areas' behavior, as a whole, is possible.

3. CLOUD VIDEO SURVEILLANCE AIDED BY GEOVISUALIZATION

Cloud computing (Armbrust et al., 2010) aims at the design of software modules that use accessible Internet on-demand services. These services are available from anywhere, with the cloud as a point of access for the consumers to all the computing resources. In Figure 1 we
show the general scheme of the architecture of our smart security approach. This architecture considers three main modules. The first one is the Smart Security module, in which all the data from the sensors, e.g. video cameras, is processed. In a second module we have the GIS, in this module we manage all the information concerning to the map and geolocalization of every object within the map, and finally in a third module, we have the web-browser. This block can be seen as the user interface where it is possible to interact with the map elements, select the type of information that will be displayed and make several interactions over the resulting visualization.

Figure 1. General architecture for a smart security in smart cities assisted by geovisualization.

3.1 Smart Security Module

The smart security module is the main component of our architecture. This component is in charge of processing all data produced by video cameras.

In Figure 2 we can identify the five sub-modules of the smart security module: (i) Video Processing (VP), which is a server for media content that eases the implementation of applications that involve video streaming. In this sub-module is where all the algorithms related to the video stream processing are implemented; (ii) Context Information Manager (CIM), used to create temporal entities capable of handling information useful for other applications, this information is available to every user subscribed to such entities. This sub-module can be seen as the moderator that carries out the communication process between the other sub-modules. (iii) Security sub-module which is required considering the context of our system (surveillance scenario), thus, the access control becomes an important element that needs to be implemented. (iv) Storage, a sub-module for data persistence capable to store relevant information in order to retrieve it later, and finally (v) Visualization, where we deal with all the information obtained from processing the video stream (video processing sub-module) which is shown to the users. For more details of how this visualization sub-module handles all the information see Section 3.2.

It is worth mentioning that all these sub-modules are based on the FIWARE 1 middleware platform, which provides a set of APIs that facilitate the design and implementation of smart applications. See (Valentin et al., 2017) in this same volume for more information about the architecture.

3.2 Visualization

In video surveillance it is common to have a set of cameras that generate a lot of information which becomes difficult to analyze by the human eye. To address this problem we have developed video processing algorithms in order to extract relevant information. This information is processed in such a way that the system is capable of alerting the user about risk situations. Furthermore, in order to make more comfortable and understandable for the user this information; is geovisualized. It is important to add that the system was designed to detect the presence of moving objects and then classify them into three classes: people, vehicles and others. In order to be able to classify any of these objects, we developed three sub-modules: background subtraction, classification and tracking. Figure 3 shows how this modules are connected. First, with the background subtraction sub-module and based on the SuBSENSE algorithm (St-Charles et al., 2015), we extract the objects that do not belong to the static part of the scene. After that, these objects are fed into the classification sub-module where a nearest neighbor algorithm (Tsoumakas and Katakis, 2006) is used to classify into one of the three predefined categories. As final step, the tracking sub-module (Lucas et al., 1981) ensures that once an object is classified and its presence is reported to the system, this object keeps its label and is only reported once as long as it stays in the scene.

Once the VP module extracts relevant information from the scene, we send this information to the CIM sub-module, which resend it to the visualization module where the graphics to be visualized are generated. In addition to this real time information, it is also possible to visualize the historical data stored at the Storage sub-module.

The resulting visualization can be of two different types: iconic or geometric. For iconic representations (pie-chart and bar-chart) a new map layer is generated and each visualization is drawn at its corresponding geographic location. On the other hand, the geometric representation is

---

1 https://www.fiware.org
visualized in an independent window to avoid overlapping. In Figure 5 is shown the way the system displays the information to the user through iconic representations and all the plotted data is placed over each camera’s location, while Figures 6 and 7 illustrate the geometric visualization technique over independent windows.

The internal architecture of the Visualization module is composed by four different sub-modules: (i) Manager, (ii) Multiparametric, (iii) Charts, and (iv) Video.

(i) Manager sub-module: this sub-module acts like a bridge for communicating the outside of the module with internal sub-modules, retrieves data from the storage module and communicates with the selected visualization technique for generating the requested visualization.

(ii) Multiparametric sub-module: This module is able to visualize multiparametric data and must be employed to visualize the global behavior of the entire smart security system at any given time.

(iii) Chart sub-module: This module allows to visualize volatile information managed by the context in real time with high performance and without losing descriptive information.

(iv) Video sub-module: Communicates the video streaming with the GIS to perform video rendering over the map.

To develop a Geovisualization-based system is important to take into account the dimensions and the shape of every resulting visualization in order to make a useful and understandable integration. Pixel-based, Icon-based, and Chart-based techniques generate iconic representations, this representation can be drawn over specific locations in a map (see Section 3.3). Conversely, the dimensionality of Geometric-based techniques impede its directly visualization over the map, but can be used for statistical exploration or for projection pursuit (Huber, 1985, Dykes et al., 2005). The use of the correct visualization method might be the difference between extracting valuable information or keep digging out in a sea of data for an undefined amount of time. Although there is no rule on how to analyze a bundle of information, certainly there are some parameters to take into account before generating a visualization for some given data.

3.3 Data visualization techniques

The general aim of the visualization techniques is to produce a representation in which it is possible to observe the correlations between the variables (Eick, 2000) without previous knowledge about the data. For the smart security scenario we are using several visualization techniques to handle data and its relation with the geospatial component. By using these techniques it is possible to provide to the users relevant information of the monitored areas.

Multiparametric data visualization techniques

Multiparametric data, also known as multidimensional data or n-dimensional data, refers to data with more than one instance and each instance with more than three variables. This family of data can be seen like a $n \times m$ matrix where $n$ and $m$ are the number of instances and the number of attributes/variables per instance, respectively. The multiparametric data visualization techniques are classified in three groups, depending on the way that graphics are deployed: (i) Icon-based, (ii) Geometry-based and (iii) Pixel-based.

Icon-based techniques have two distinguishable parameters: the type of geometry for representing the instance and the position defined for every icon in the image. These techniques are recommended when the number of dimensions/variables is between ten and fifteen, and the number of instances is large. Examples of these are Chernoff Face, Start Fields, Shape Coding, etc. (Chernoff, 1973, Eick, 2000). On the other hand, Geometry-based use geometrical entities as visual properties to produce visualizations (lines, curves, points). Among the set of existing techniques, there are three which are the most used: parallel coordinates, dispersion diagram, and Andrews plot (Inselberg and Dimsdale, 1990, Andrews, 1972). Pixel-based techniques reduce the visualization space using a pixel codification for each instance’s variable. The codification is always through a color value and every instance has as many pixels as attributes/variables has. The more common used techniques are circle segments (Ankerst et al., 1996) and recursive patterns (Keim et al., 1995).

Chart-based visualization techniques

The visualization techniques based on charts, can be divided in three main groups: (i) Bar-based: are used to compare the quantities of different classes (histograms), (ii) Segment-based: separate representations in sections called slices where the amount of data is given by the arc-length of the slice (pie-charts) (Bezjian, 1996) and (iii) Line-based: are most often used to illustrate the relation among several variables (time series), each one represented by a line as in (Phillips et al., 1938).

We have implemented three visualization techniques. Two with iconic representation (chart-based) and one geometrical (multiparametric). The two Chart-based (pie-chart and bar-chart) techniques implemented in the system can be used for both, visualize the volatile information which is arriving to the context from the VP and for visualizing small time windows of stored data. The multiparametric technique implemented (parallel coordinates) is only used to visualize large amounts of historical data.

4. CASE STUDY

To test the smart security system we use as test scenario the INAOE’s campus. To visualize the map we use Openstreetmap and for data visualization we used D3.js (parallel coordinates visualization) and ExtJS (charts visualization and GUI) javascript libraries. All the data is coming from a video surveillance system prototype that we already have implemented. This prototype consists of four geolocalized IP cameras. In order to get latitude and longitude of each camera we used the GPS of a smartphone.

Cameras were placed in a real scenario where it is difficult to have control of the environment conditions, this represents a challenge for the computer vision algorithms (e.g.}

2https://www.openstreetmap.org/
background subtraction, tracking and person recognition). In Figure 4 we depict the map of our campus as well as the location of the cameras and its respective fields of view.

As we previously mentioned, all the blocks of our proposed smart security module are based on FIWARE. This middleware provides a set of functional components, known as generic enablers, to develop applications for Future Internet. Specifically, we are using four of these components: the Orion Context Broker (OCB), Kurento, Cygnus and Cosmos.

In order to show how our visualization module works lets see the following example. Suppose the user wants to see the amount of people, vehicles and other objects that transited in front of each of the four cameras from “may 8 13:00:00 CDT 2017” to “may 8 15:00:00 CDT 2017”. To do that, we perform the communication between modules through json messages. In Listing 1 we are showing the type of message to make a query and then generate a visualization such as the one depicted in Figure 5 (a).

**Listing 1 JSON example**

```json
{
  "startTime": "may 8 13:00:00 CDT 2017",
  "endTime": "may 8 15:00:00 CDT 2017",
  "camerasId": ['a', 'b', 'c', 'd'],
  "visTech": "charts:piechart"
}
```

If the user wants to change the visualized icon, the name of the technique must be specified in the parameter visTech like in Listing 2 (Figure 5 (c) shows the resulting visualization).

**Listing 2 JSON example**

```json
{
  "startTime": "may 8 13:00:00 CDT 2017",
  "endTime": "may 8 15:00:00 CDT 2017",
  "camerasId": ['a', 'b', 'c', 'd'],
  "visTech": "charts:barchart"
}
```

Besides, if an event is detected, the user has the option to see the video stream in real time just clicking the camera icon. The message generated to retrieve video data is depicted in Listing 3 and results are shown in Figure 5 (b).

**Listing 3 JSON example**

```json
{
  "camerasId": ['d'],
  "visTech": "video:video"
}
```

For the sake of simplicity, we will first explain how to interpret graphics generated with the parallel coordinate technique for only one video stream. In Figure 6 we are showing an instance of this visualization technique. The first axis represents the hour of the day (in hours). To identify the hour of the day we are using lines with different colors (from green to red). The second axis is used to represent the camera id. The remaining axes are used to represent the attributes (persons, vehicles, others) that we want to observe. This graph is interpreted as follows, the intersections between each horizontal line with each of the vertical axes denotes the value of the attribute for a specific hour of the day.

Figure 6 shows that the camera (a) was working for the entire day, during this period of time the number of persons never exceeds 40. Furthermore, we can observe that cars and other objects are not detected in this monitored area. For this figure we are visualizing the content of only one camera, however, all the cameras might be visualized at the same time (See Figure 7).

5. CONCLUSIONS AND FUTURE WORK

In this paper we have proposed a novel visualization approach for smart video surveillance that integrates web-based Geographic Information Systems and Scientific Visualization techniques. Our approach gives the user the possibility of visualizing large amounts of data through different layers of information, which in turn allows him to have a better idea of the global behavior of the monitored areas. Besides, this system gives the user the option of watching the video streaming of a specific area, so that the user can take decisions related to an alert. In addition, the proposed approach allows to incorporate another kind of sensors and visualize the data generated by these sensors. We have implemented our system in a real scenario using the FIWARE middleware. As a future work we want to extend the set of algorithms in the VP sub-module and to evaluate the visualization techniques with experts in the video surveillance area.

ACKNOWLEDGEMENTS

This work was supported in part by FONCICYT (CONACYT and European Union) Project SmartSDK - No. 272727. Reineri Oyes García is supported by a CONACYT Scholarship No.789638

---

3https://catalogue.fiware.org/enablers/publishsubscribe-context-broker-orion-context-broker
4http://www.kurento.org/
Figure 4. Map generated by OpenLayers of INAOE’s facilities. Red marks represent the geolocation of each of the 4 cameras. On the right images we have (in cyan) the field of view of each camera (best seen in color).

Figure 5. Resulting visualization of the number of persons (green), vehicles (blue), and others (red) detected by each camera during two hours (see queries 1, 2, and 3). Left figure shows a Pie-chart visualization. Center figure shows a Pie-chart visualization and real time video rendering of camera d). Right figure depicts a Bar-chart visualization. In this figure is shown that camera a) does not recognize vehicles because it is in an indoor scenario. Cameras c) and d) have a considerable amount of persons because it correspond to the lunch time. Finally, b) camera has a lot of vehicles because it is in the parking lot (best seen in color).
REFERENCES


GIS, G., 2011. Geographic resources analysis support system.


Vázquez R., R., 2016. Nuevos métodos para el procesamiento y análisis de información geográfica.