

# An integrated system of video surveillance and GIS

Jin Zheng<sup>1, \*</sup>, Dongni Zhang<sup>1</sup>, Zhi Zhang<sup>1</sup> and Xiaojun Lu<sup>2</sup>

<sup>1</sup>School of Computer Science and Engineering, Beihang University, Beijing, China

<sup>2</sup>China International Engineering Consulting Corporation, Beijing, China

\*Corresponding author e-mail: JinZheng@buaa.edu.cn

**Abstract.** Nowadays a great number of cameras have been installed in cities and generate tremendous video data every day. To manage the fragmented video data generated by these independent cameras, this paper designs and implements an integrated system of video surveillance and GIS. Using the advantages of GIS in geographical space, the system can manage the tremendous cameras and video data from both space and time aspects. In terms of space, the spatial distribution and monitoring range of the cameras can be displayed on the GIS map, and the layout density of the cameras can be provided. More importantly, combining the road network and the locations of cameras, the system can obtain the accurate topological relationship between cameras, which includes the reachability of the road network and the actual distance between cameras. In terms of time, the transition time between cameras can be acquired based on the topological relationship, and the spatial-temporal relationship between the various videos of different cameras can be established. The experimental results prove that the system can manage the independent cameras and the fragmented video data effectively.

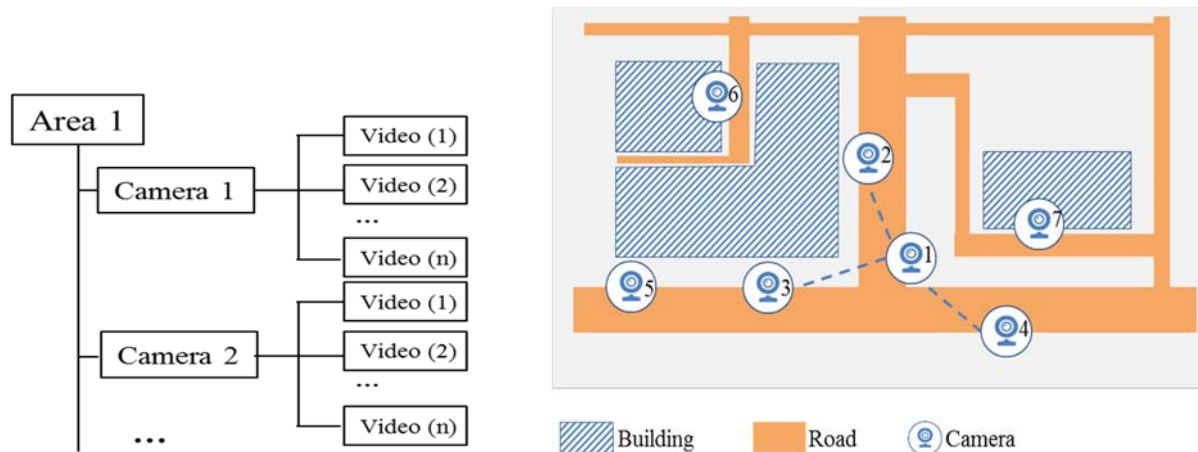
## 1. Introduction

With the requirement of the construction of safe city, and the reduction of the cost of video surveillance equipment, many cities have installed a large number of security cameras. Many video surveillance systems have also been established [1], which play significant roles in various fields, such as urban traffic, public security and military reconnaissance [2]. The number of cameras is increasing, resulting in tremendous video data. To manage the tremendous cameras and video data, the traditional methods establish a tree structure according to the monitoring area of the cameras (shown in Fig.1). Each video sequence is attached to a camera according to the ownership, which means the video sequence is captured by the camera. However, the locations of the cameras are not intuitive, and the reachability between cameras is unclear. The video data captured by different cameras is isolated and separated, which results in the fragmentation of the information collected by the cameras.

In order to effectively manage the cameras and fragmented video data, many researchers focus on the topological relationship of cameras [3-6]. Traditionally, the topology network of cameras is set up by manual measurement. Fig.2 shows the diagram of a network topology established by manual measurement. Each camera has an attribute, which describes the reachability among cameras. Obviously, this method requires that the user is familiar with the installation environment, and is time-consuming due to a large number of cameras. Thus, the research of obtaining the camera topology automatically has attracted wide attention [3]. Nam presents an inference approach to build the camera network topology [4]. T. et.al propose an algorithm of estimating the network topology of cameras based on



weighted time window [5]. Chen proposes an unsupervised learning approach to estimate the transition time between cameras [6]. In general, these approaches identify the time that objects appear and disappear in the camera views based on visual tracking and re-identification. Although these approaches only deal with video data to establish the topology network of cameras, they are difficult to handle the complex scene with massive cameras. Moreover, the approaches are limited in application for the serious time consuming. Obtaining the topological relations of cameras is still facing the challenges of automation, accuracy and speed.



**Figure 1.** Tree structure of data management **Figure 2.** The diagram of camera topology

Actually, the obtainment of topological relations is inadequate to meet the needs of effective management of cameras and video data, because the effective management also involves the display of the camera location and monitoring range, as well as the spatial-temporal relationship of various video data. GIS (Geographic Information System) technology has potential abilities to breakthrough this challenge and solves these problems. Firstly, GIS is able to obtain the spatial position of cameras directly and accurately, and it has the advantage of processing spatial information efficiently [7, 8]. Secondly, the powerful visualization function of GIS can display the spatial distribution of cameras clearly. Moreover, GIS can provide the road network for the judgment of reachability. Therefore, the combination of GIS and video surveillance has been widely studied. For example, Zhang proposes a cross-mapping approach of video and 2D GIS [9]. Song applies GIS in the real-time monitoring of crowd counting, which can display the crowd counting data on the map and provide the warning information [10]. However, these researches only make the content of the video mapped onto a corresponding point of an electronic map according to spatial position, and the relationship among these points is still not been extracted. Thus, the existing integrated systems of video surveillance and GIS are far from the efficient management of cameras and video data.

This paper designs and implements an integrated system of video surveillance and GIS, which fuses the road network information and camera position to obtain the topological relations of cameras. The integrated system not only display the spatial distribution, the monitoring range and the layout density of the cameras on the GIS map, but also establish the spatial-temporal relationship between the videos of various cameras. The experimental results reveal the validity of the system.

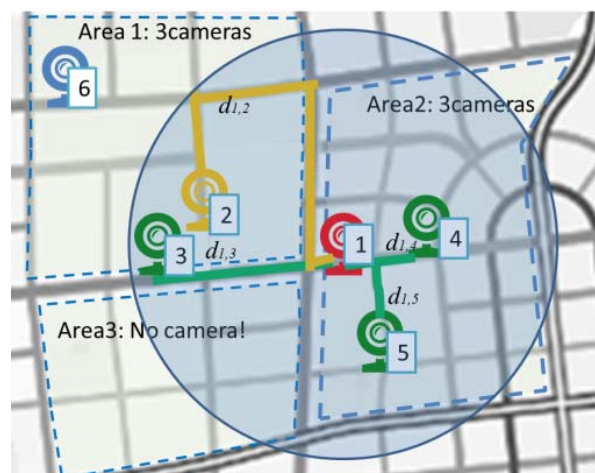
## 2. System design

### 2.1. Function design

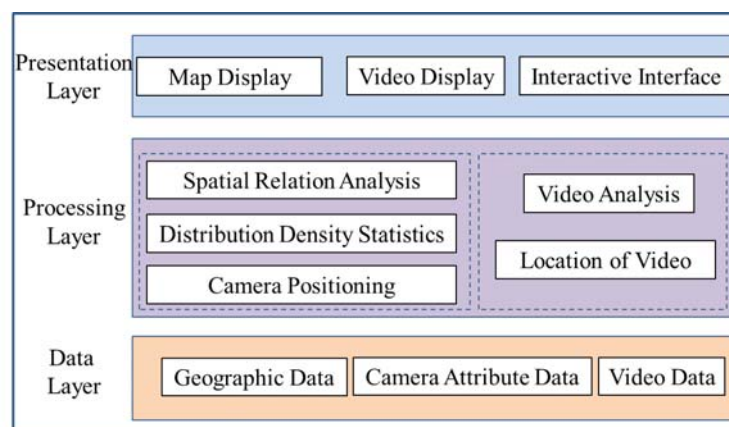
As the purpose is to provide the management and display of cameras, as well as establish the spatial-temporal relationship of video data, the function design of the integrated system mainly focuses on two aspects.

**Spatial aspect:** (1) each camera can be displayed accurately as a marker on the map. (2) The monitoring range of each camera can be displayed. (3) The number of cameras in a certain area can be counted and displayed, and the monitoring blind area is also been pointed out. (4) The spatial reachability between cameras can be expressed. The distance between two points is computed along the road, not the Euclidean distance on the map. (5) Based on the spatial reachability, the system is able to search the nearby cameras of the selected camera automatically and provide the actual distances between the selected camera and its nearby cameras. The details are shown in Fig.3.

**Temporal aspect:** (1) according to the road network, the moving time of object from one camera to another camera can be calculated, which is called transition time. (2) Based on transition time, the temporal relationship of video data captured by different cameras can be established. (3) The users can extract the associated video clips which are from various cameras, and achieve continuous object tracking among multiple cameras.



**Figure 3.** Schematic diagram of functional design



**Figure 4.** Software architecture of the system.

## 2.2. Architecture of the system

The integrated system uses three-layer architecture, which contains data layer, processing layer and presentation layer, as depicted in Fig.4.

For data layer, geographic data is the foundation of GIS functions. Camera attribute data contains the longitude and latitude data of cameras, as well as the monitoring distance. Both geographic data and camera attribute data support the process of camera positioning, distribution density statistics and spatial relation analysis. Video data is captured by cameras and is the base of video analysis.

For processing layer, GIS and video surveillance are two core modules. Camera positioning is the basic mapping process, and distribution density statistics is to count the number of the cameras installed in a certain area. Spatial relation analysis mainly includes searching the reachable cameras and calculating transition time. Video analysis includes object detection, object tracking and object re-identification, which is mainly used for video object processing in single camera. Location of video is to find the video clips which are captured by other cameras but associated with the current video.

For presentation layer, as an interface between the system and users, it consists of map, markers display, video display and interactive interface. It is used to accept the request and return the response results.

### 3. System implementation

#### 3.1. The work flow of spatial-temporal relationship analysis

Although the integrated system has many functions, such as map and camera positioning, distribution density statistics, and markers display etc., the core of the system is to establish the spatial-temporal relationship between the videos of different cameras, which can be used for the management of cameras and video data. In the integrated system, a typical application of the spatial-temporal relationship is multi-camera object tracking. Thus, multi-camera object tracking is used to illustrate the application of spatial-temporal relationship in the integrated system. The processing work flow is as follows.

Step 1: object tracking in single-camera is done and object features are contained.

Step 2: if the tracked object disappears in a camera view, the system returns the attributes of its reachable cameras, including the actual distances and transition times.

Step 3: According to the results of Step 2, a set of video clips are extracted out from the reachable cameras.

Step 4: Based on the video clips, object re-identification is executed.

Step 5: The video clips containing the tracked object are sorted according to the appearing time of the object.

Step 6: The trajectory of the tracked object is tagged on the GIS map.

In this way, the management of cameras and video data is unified as the maintenance of spatial-temporal relationship of video data.

#### 3.2. The implementation of spatial-temporal relationship analysis

The functions of GIS are supported by the online map service. We call the AMAP API based on Web technology [11] to achieve the basic map operations and geographic analysis. In addition, aiming at the establishment of spatial-temporal relationship, the executed statements are list in Table 1.

$T\_search$  means how long the tracked object has left from the current camera view.  $V$  means the average moving speed of the tracked object, which is computed based on object tracking in single-camera. According to  $T\_search$  and  $V$ ,  $R\_search$  is obtained to expresses the maximum Euclidean distance of the object movement in a given time  $T\_search$ .

$P\_cur$  is the position of the current camera. According to  $P\_cur$  and  $R\_search$ , a candidate camera set  $Cam\_a \{C_1, C_2, \dots, C_i, \dots, C_n\}$  is obtained through the Search Nearby function, which is provided by the online map service.  $P\_cur$  and  $R\_search$  are two parameters of the function, which mean the center point and searching radius. The candidate camera set  $Cam\_a$  includes all the cameras located in the generated circle.

*Actual Distance*  $\{D_1, \dots, D_n\}$  between  $P\_cur$  and each element of  $Cam\_a$  is computed by the shortest path analysis function ShortestPathPlanning based on GIS service. According to *Actual Distance* and  $V$ , *Transition\_time* of each element in  $Cam\_a$  can be calculated.

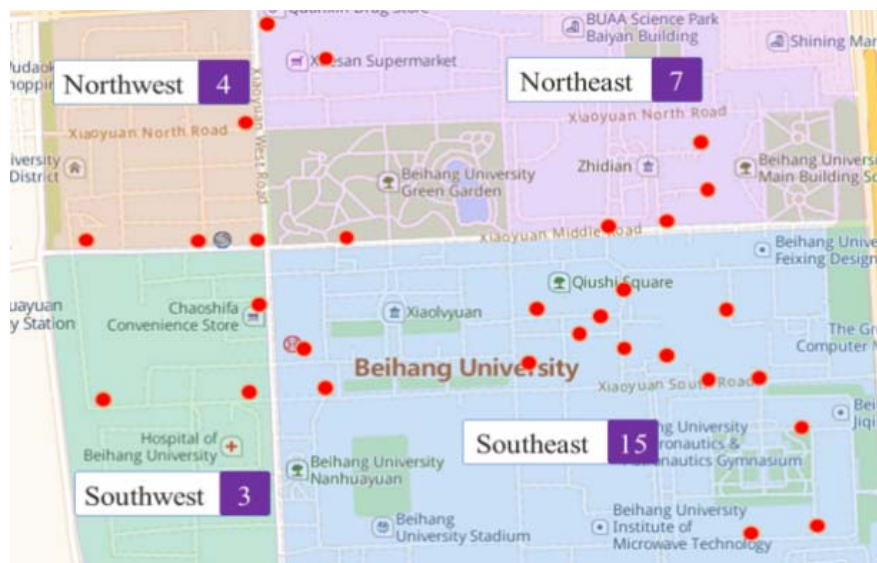
If the condition that  $Transition\_time < T\_search$  is satisfied, the element in the set  $Cam\_a$  is picked up. The condition indicates that the tracked object is possible to reach the camera position within the time of  $T\_Search$ . The elements which satisfy the condition form a new set  $Cam\_b$  as an output, and their attributes can be marked on the map.

**Table 1.** The executed statements of spatial-temporal relationship analysis.

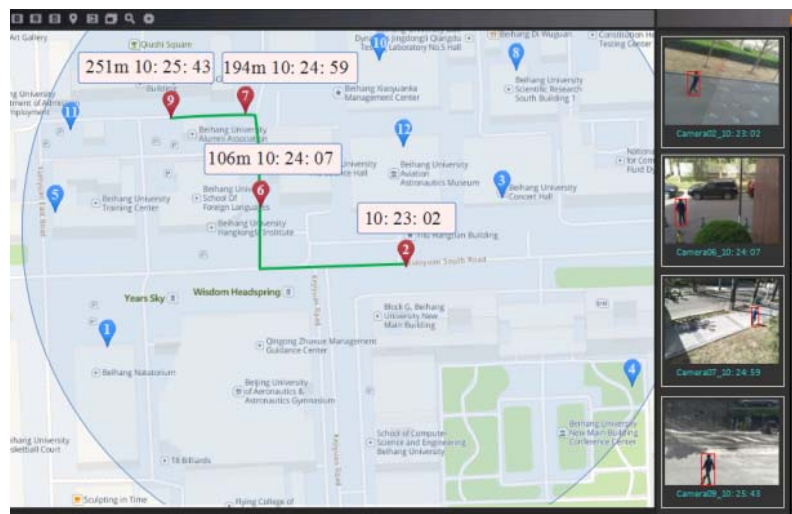
Input	Function	Output
$T\_search, V$	$R\_search = T\_search \times V$	$R\_search$
$P\_cur, R\_search$	Search Nearby ( $P\_cur, R\_search$ )	$Cam\_a\{C_1, \dots, C_n\}$
$P\_cur, Cam\_a\{C_1, \dots, C_n\}$	ShortestPathPlanning ( $P\_cur, \{C_1, \dots, C_n\}$ )	$ActualDistance\{D_1, \dots, D_n\}$
$ActualDistance\{D_1, \dots, D_n\}, V$	$Transition\_time = ActualDistance/V$	$Transition\_time$
$Cam\_a\{C_1, \dots, C_n\}, Transition\_time$	If ( $Transition\_time < T\_search$ )	$Cam\_b\{C'_1, \dots, C'_m\}$

#### 4. Experimental results and application

Taking the integrated system installed in our campus as an example, the effectiveness of the system is illustrated. As is shown in Fig.5, the red markers represent the installed cameras, and the statistical density in each area is also displayed on the map. Fig. 6 shows an example of object tracking in multi-camera. When the object disappears in a camera, the system automatically finds the candidate cameras (located in the circle area). According to the actual path and transition time, the reachable cameras (drawn with the red marker) can be picked up. Furthermore, the video clips (see the frames on the right) related to the same object but from different cameras are organized with the order of appearance time. As a result of multi-camera tracking, the moving trajectory is displayed on the map (see the green curve). Compared with the methods of estimating the network topology of cameras by training data, the proposed system obtains the accurate topological relations of cameras fast. In addition, the graphical depiction in GIS also makes the results easier to be learned and displayed.

**Figure 5.** Map and marker display





**Figure 6.** An example of object tracking

## 5. Conclusion

For the effective management of massive cameras and video data in video surveillance, it is crucial to obtain the spatial-temporal relationship of cameras fast and accurately. We design a framework of integrated system combining GIS and video surveillance, which has the ability to acquire the accurate spatial-temporal topological relations of cameras. Thus, the isolated cameras and the fragmented video can be effectively managed based on the spatial-temporal topological relations of cameras. The experimental results prove the validity of the proposed integrated system.

## Acknowledgments

This work is supported by the National Key Research and Development Plan (No.2016YFC0801002), the NSFC (No.61370124, No.61632001, No. 61772054).

## References

- [1] Shao Z, Cai J, Wang Z, et al, Analytical Processing Method of Big Surveillance Video Data Based on Smart Monitoring Cameras, *Journal of Electronics and Information Technology*, 2017.
- [2] Shao Z, Cai J, Wang Z., Smart Monitoring Cameras Driven Intelligent Processing to Big Surveillance Video Data, *IEEE Transactions on Big Data*, pp (99): 1 - 1, 2017.
- [3] Wang Y, Lu K, Zhai R, et al, Challenge of Multi-camera Tracking, *International Congress on Image and Signal Processing*, pp. 32 - 37, 2014.
- [4] Nam Y, Rho S, Park J H, Inference Topology of Distributed Camera Networks with Multiple Cameras, *Multimedia Tools and Applications*, vol. 67, No. 1, pp. 289-309, 2013.
- [5] T. Feigang, H. Ling, L. Jian, Pedestrian Tracking across Non-overlapping Cameras Based on Fusion Multi-clue, *Application Research of Computers*, vol. 34, No. 8, pp. 2542-2545, 2017 (In Chinese).
- [6] Chen X, Huang K, Tan T, Learning the Three Factors of a Non-overlapping Multi-camera Network Topology, *Pattern Recognition*, Springer Berlin Heidelberg, pp. 104 - 112, 2012.
- [7] M. Aleksandar, D. Aleksandar, R. Dejan, GIS-augmented Video Surveillance, *International Journal of Geographical Information Science*, vol. 24, No. 9, pp. 1415 - 1433, 2010.
- [8] M. Aleksandar, R. Dejan, D. Aleksandar, et al, Integration of GIS and Video surveillance, *International Journal of Geographical Information Science*, vol. 30, No. 10, pp. 2089 - 2107, 2016.
- [9] Z. Xingguo, L. Xuejun, S. Hongquan, Video Surveillance GIS: A Novel Application, *Proceedings of the 21st International Conference on Geoinformatics*, Kai Feng: IEEE, 2013.
- [10] S. Hongquan, L. Xuejun, Z. Xingguo, H. Jiapei, Real-time monitoring for crowd counting using

video surveillance and GIS, Proceedings of the 2nd International Conference on Remote Sensing, IEEE, pp. 1 - 4, 2012.

[11] Information on <http://lbs.amap.com>