A Retrieval Method Based on Video Fusion for Surveillance Systems

Cheng-Hsiung Hsieh, Hsien-Chou Liao, Chih-En Wu, and Yi-Ming Hsieh
Department of Computer Science and Information Engineering
Chaoyang University of Technology
E-mail: {chhsieh, hcliao, s9527611, s9527610} @cyut.edu.tw

Abstract

Digital surveillance systems have already been used in community security widely. However, one main drawback of current surveillance systems is that the user has to search videos from different cameras when an event happened. For example, when tracking an object, one needs to inspect videos from different cameras which spends much time generally. In this paper, we propose a video fusion and retrieval method to solve the problem. By fusion and retrieval techniques, we are able to track a specific moving object automatically. This makes surveillance systems more convenient. A community scenario is given to demonstrate the proposed method.

Keywords: *community surveillance systems, video fusion, video retrieval*

1. Introduction

As the advance of technology, surveillance systems are getting more powerful and cheap. Consequently, they are extensively used in different places where the security is concerned. One example is to equip a surveillance system in a community. Currently, functions in a surveillance system are quite simple. They are recording and moving object detection in general. Recently, several approaches have been reported to improve the performance and application of surveillance system. In [4], a surveillance system was proposed to adapt with different situations from parking areas, commercial districts to housing. In [3], an intelligent visual surveillance system was presented whose function having intruder classification. In [2], a video-based automated event detection framework was included to perform object segmentation, multiple objects tracking, object trajectories analysis, and relevant events reporting. In [6], a 3-D particle filter based object tracking was reported. It was able to predict and update the probability distribution of moving objects in 3-D domain for multi-camera surveillance system. In [5], with sensors a security system of multi-camera was proposed. In [1], an embedded system with smart cameras was presented which combining video sensing, video processing, and communication within a single device.

Though those approaches have improved the performance of surveillance system to some extent, it, however, is not convenient for the user to review videos when tracking a suspicious object. Generally, in the case the user finds the object by time and location with videos recorded by different cameras manually. This usually takes a lot of time. To deal with the problem, in this paper we propose a video fusion and retrieval method, which is abbreviated as VFR method. This paper is organized as follows. Section 2 describes the proposed VFR method. Then a scenario is provided to demonstrate the VFR method. Finally, conclusion and further research are given.

2. The Proposed VFR Method

To facilitate a fast and automatic tracking a moving object, a video fusion and retrieval method is incorporated into a surveillance system. Basically, the proposed VFR method consists of three parts: moving object detection and information extraction, the user query, and video fusion. These parts are described in the following.

2.1 Moving Object Detection and Information Extraction

For the moving object detection, a common used scheme background subtraction is applied where background frame is updated for every 25 frames for better detection. A screenshot for moving object detection is illustrated in Figure 1. With camera number and corresponding filename, four attributes of detected object are recorded for every second: current time, upper left corner coordinates, length, and width of moving object. Those extracted information are recorded in

XML format because of its convenience. One XML file to record the information of moving object is shown in Figure 2.

2.2 The User Query

For a moving object of interest in a certain camera, the user may input time and location to retrieve corresponding videos. With the time and location information, video clips in different cameras are fused according to the time stamps in corresponding XML files. The video fusion process is described in the following section.

2.3 Video Fusion

Given a moving object, the information recorded in XML file is analyzed and then used to fuse related videos. In the given camera of the selected moving object, a reference frame is chosen and the upper left corner coordinate is considered as reference point. Then the up-coming frames of moving object are compared with the reference point. By this doing, the moving direction can be inferred from which the following camera number with the moving object can be determined as well. The process described above continues for the selected camera. The iterative process terminates when the moving object leaving the scope of surveillance system. The compound information from different cameras is linked through which all related video clips are fused. The fused video is the response of user query.

3. VFR-Based Surveillance System

In this section, a system with fusion-based retrieval mechanism is presented. Note that current surveillance systems generally have no operation interface for users. The user must search in many videos manually which usually need to spend much time. Also, the user needs to track an object of interest in different cameras. This is a pain for the user. Consequently, the system provides an operation interface to automatically track the moving object. The system structure is given in Figure 3 which is composed of three parts:

(A) Preprocessing
In this phase, moving objects in all cameras are detected where corresponding attributes are extracted and recorded in XML files.

(B) Operation interface
The interface provides user to input query information and display the fused video which is requested by the user.

(C) Fusion process
According to the user query, related video clips are fused by the information recorded in XML files. The final output is displayed in the operation interface.

4. Scenario in Community Surveillance System

The proposed system described in Section 3 is implemented by Visual Studio.NET and run on a PC. Test video files are from 14 cameras in a real-world community surveillance system, whose time spanned from 7 to 8 o’clock in the morning of some day.

By the method stated in Section 2.1, the moving object is detected. Figure 1 is an example where the red rectangle shows the located moving object. The information of the moving object is extracted and recorded in a XML file as shown in Figure 2.

The operation interface is shown in Figure 4 which includes the query input, searching function of a moving object, the fused video display, the community map with camera allocation, and the trajectory of the moving object.

The fusion process for related video clips is given in Figure 5 where an example is demonstrated. And the video clips for different cameras are shown as well.

In the scenario, the VFR-based system performs well for tracking a moving object most of time. This increases the convenience in searching an interested object when needed.

5. Conclusion and Future Work

In this paper, we proposed a video fusion method to deal with the inconvenience of tracking a moving object in current surveillance system. A prototype VFR-based surveillance system is constructed and applied to a community surveillance system. The experimental result showed that the proposed system worked well in the scenario most of time. However, there are some aspects should be considered to improve the prototype such as amending attributes for more accurate tracking, multiple objects tracking, and moving object recognition. These considerations will put into account for the VFR-based surveillance system in the further research.

References

[1] Bramberger, M., Doblander, A., Maier, A., Rinner, B. and Schwabach, H., "Distributed Embedded Smart Cameras for Surveillance Applications," *IEEE Computer Society*, Vol 39, Issue 2, pp. 68- 75, Feb. 2006.

[2] Cheng Hsu-Yung and Hwang Jenq-Neng, "Automated Event Detection for Video Surveillance via Hidden Markov Models," *Computer Vision, Graphics and Image Processing Conference*, pp. 165-172, 2007.

[3] Hu Wu Chih, Su Chung-Yuan, Chuang Chi-Chihm and Huang Chun-Wei, "Intelligent Visual Surveillance in Net Cage," *Computer Vision, Graphics and Image Processing Conference*, pp. 158-164, 2007.

[4] Thiel G., "Automatic CCTV Surveillance Towards the VIRTUAL GUARD," *IEEE Aerospace and Electronic Systems Magazine*, Vol 15, Issue 7, pp. 3-9, Jul. 2000.

[5] Trivedi, M.M., Gandhi, T.L. and Huang, K.S., "Distributed Interactive Video Arrays for Event Capture and Enhanced Situational Awareness," *IEEE Intelligent Systems*, Vol 20, Issue 5, pp. 58 -66, Sep. 2005.

[6] Wang Yu-Ching and Wang Sheng-Jyh, "Moving Objects Tracking and Correspondence over Multi-Camera Surveillance System," *Workshop on Consumer Electronics and Signal Processing*, pp. 47-52, 2007.

Figure 1. A screenshot of moving object detection


```
<Cameras>
    <CameraNumber>03</CameraNumber>
    <Filename>03 (04-09-06 07&apos;00&apos;38).xml</Filename>
    <Time>20060904070038</Time>
    <Object>
        <Time>20060904070038</Time>
        <Left>40</Left>
        <Top>72</Top>
        <Width>71</Width>
        <Height>167</Height>
    </Object>
    <Object>
        <Time>20060904070039</Time>
        <Left>40</Left>
        <Top>72</Top>
        <Width>71</Width>
        <Height>167</Height>
    </Object>
```


Figure 2. Information recorded in XML file for a moving object

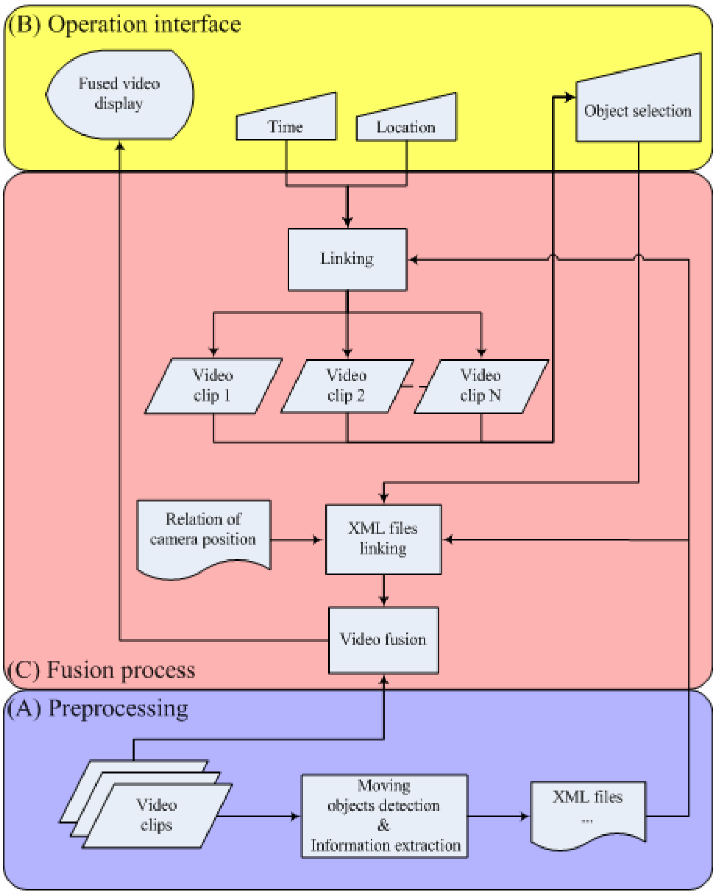

Figure 3. The structure of VFR-based surveillance system

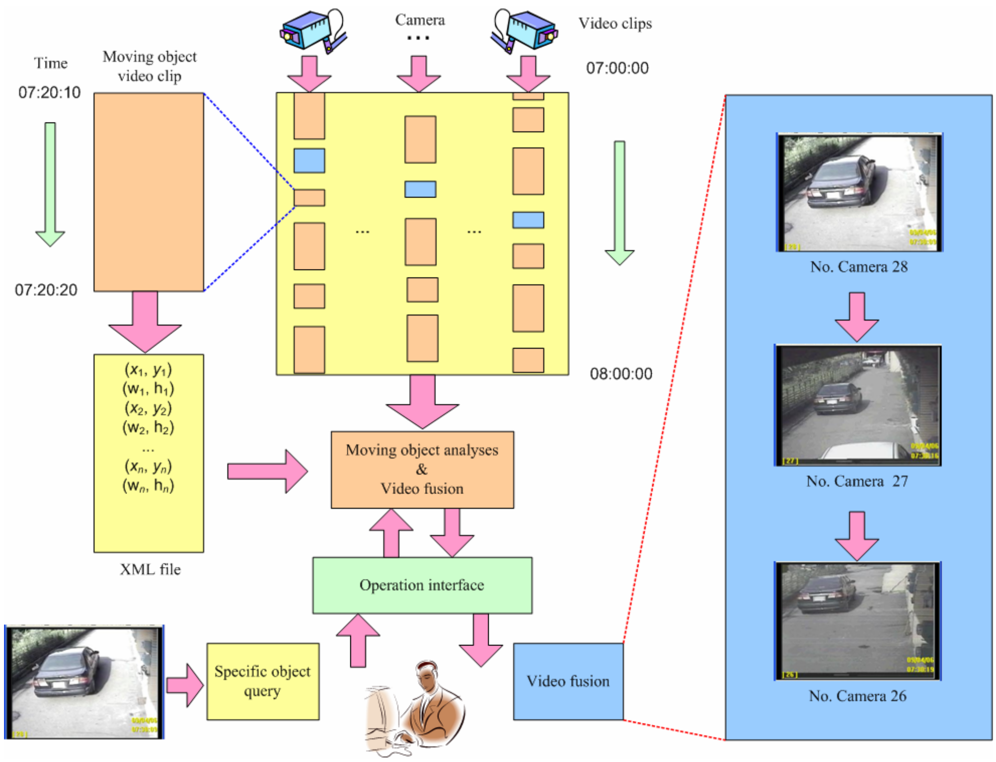

Figure 4. An example of video fusion process

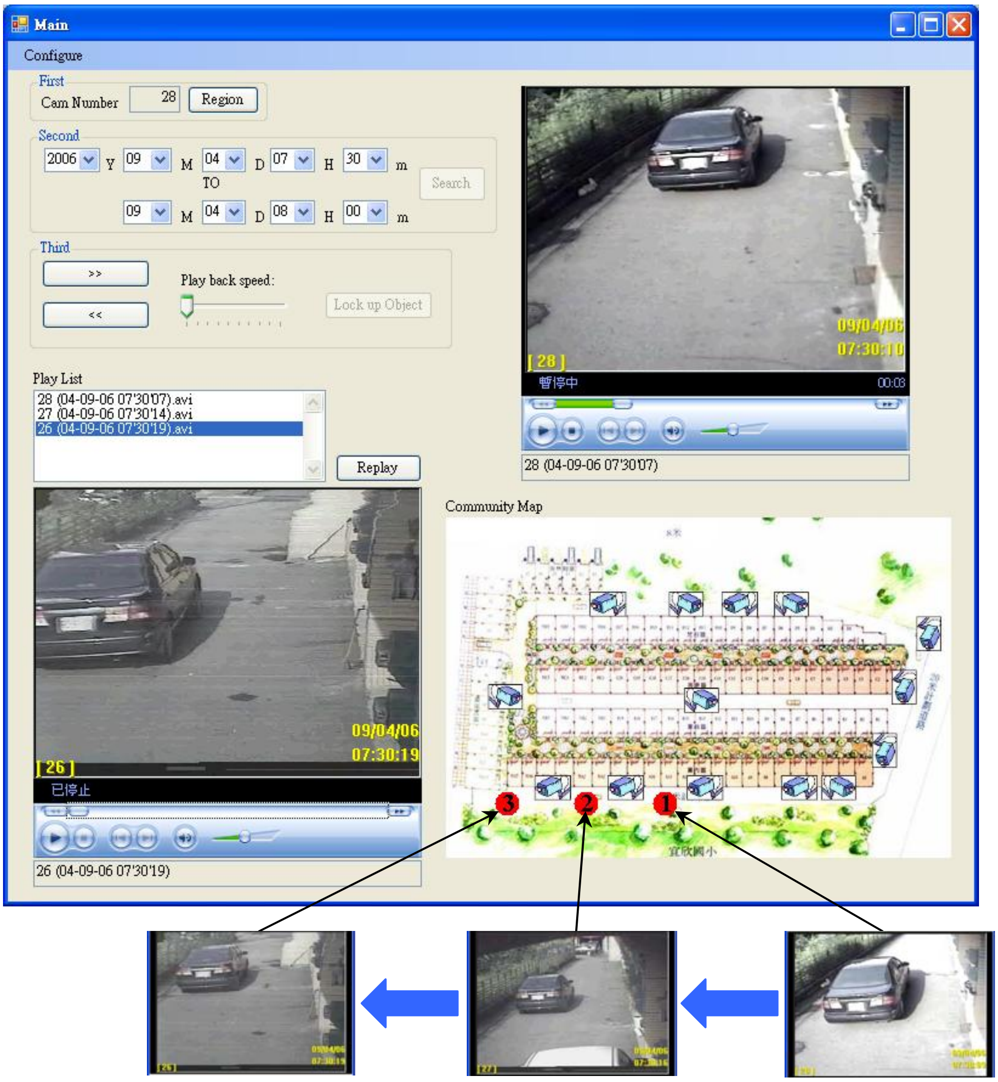

Figure 5. The operation interface with an example