Eighth International Conference on Intelligent Systems Design and Applications

A Method of Camera Relationship Establishment Based on Temporal and Spatial Information of Video Clips

Hsien-Chou Liao, Cheng-Hsiung Hsieh, Yi-Ming Hsieh, and Chih-En Wu
Department of Comp. Sci. and Info. Eng., Chaoyang University of Technology
No. 168, Jifong E. Rd., Wufong Township, Taichung County, 41349, Taiwan (R.O.C.)
Tel: +886-4-23323000~4211 Fax: +886-4-23742375
E-mail: {hcliao,chhsieh}@ cyut.edu.tw

Abstract

Camera is widely used in daily life in recent years and causes our life towarding the ubiquitous camera environment. The analysis of camera videos offers the functions of detecting, classifying or identifying moving objects. The cooperation of multiple cameras for offering intelligent services has also become an important issue in recent studies. The cooperation relies on a pre-defined camera spatial relationship. It allows the prediction of the next possible camera when an object leaves the FOV (field-of-view) of one camera. To simplify the establishment of the camera spatial relationship, a method is proposed based on the analysis of video clips recorded from multiple cameras. The information of moving objects in the video clips is used to establish correct spatial relationship. The video clips of a community surveillance system with 13 cameras are used for the experiment study of our method. The results show that the method can establish camera spatial relationship correctly.

1. Introduction

As cameras are widely used in daily life, including community surveillance, road surveillance, and home security, many related research subjects are studied, i.e. traffic flow analysis, vehicle classification, or collision detection and so on [1-4]. The analysis of camera image is used to generate critical information for surveillant, e.g., flame detection, missing/unattended object, or duress detection. In the recent years, the study and applications of surveillance cameras focus on the cooperation of multiple cameras to achieve various intelligent services. For example, the study presented by J. Black et al. in 2006 has adopted multiple cameras to monitor the same area. The location information derived from one camera can be used to correctly resolve and identify the size of an object comprehensively even if a portion of an object is obscured from the image of another camera. In addition, the researchers O. Javed et al. have in 2003 presented the Knight system [6]. It is used to track specific moving objects by coordinating the varied surveillance area covered by multiple cameras installed in a corridor. O. Javed et al. have also in 2007 proposed another automated surveillance visual system [7] in a physical environment. Suppose all camera spatial relationship can be established automatically, it will greatly contribute to the feasibility and convenience of these methods.

Under the circumstances of a fewer number of cameras, defining the cameras' spatial relationship has not been a problem. However, when the number of cameras is large, the definition of camera relationships tends to cause system deployment problem. Besides, if there is a need to modify the relationship definition whenever cameras are added or removed, it is inconvenience to those methods and systems. Therefore, the ability to establish these relationships automatically will be useful to enhance the feasibility and convenience of these methods.

2. Methods

Suppose that we are focusing on the surveillance system at a community, where the FOVs of cameras would cover all roads in the community. Their FOVs may be either overlap or a distance apart. Each camera is assigned a unique number. The recording function is activated through motion detection. That is, the surveillance system will only begin recording the video clip when detecting a moving object. The format of video clips is AVI with 25 frames per second, and each frame is 320x240 pixels. According to the above environment, the process of the proposed method is depicted in Figure 1.

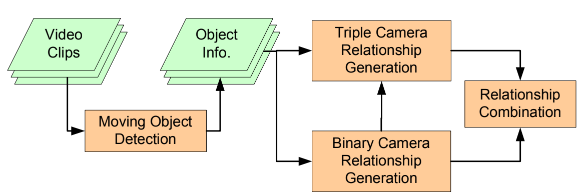


Figure 1: The process of our method

978-0-7695-3382-7/08 $25.00 © 2008 IEEE
DOI 10.1109/ISDA.2008.58

2.1 Moving object detection

A background subtraction method is used to locate moving objects [8, 9]. The first frame of each video clip is deemed as the background and updated every 25 frames to avoid the significant change in background, e.g., brightness. Figure 2 depicts an example of the detection of moving objects. In Figure 2(a), the label "[3]" indicates the video was recorded from the number 3 camera. The red rectangle depicts the detected moving object where it is moving away from the camera. In Figure 2(b), the object is moving toward the camera.

(a) (b)

Figure 2: Two examples of moving object detection

The moving object information is stored in a corresponding XML file, whose content includes the information such as camera number, file name, recording starting time, object size and related information.

2.2 Binary camera relationship generation

In this step, the direction of a moving object in a video clip is detected to determine whether the object is moving toward or away from the camera. A partial result derived from this step is listed in Table 1.

Table 1: Partial moving directions in video clips

Video clip	Direction
06 (20-02-08 07'02"04).xml	L
06 (20-02-08 07'04"57).xml	L
06 (20-02-08 07'08"27).xml	C
06 (20-02-08 07'27"37).xml	U
12 (20-F3-08 07'02"09).xml	L
18 (20-02-08 07'02"48).xml	L

The notation "C" (coming) indicates that an object is moving toward the camera, "L" (leaving) indicates that an objective is moving away from the camera, "U" (unknown) indicates it is unable to detect the direction. The possible scenarios of unknown direction include, an object is standing still, or the time an object appears is too short.

After an object's movement is detected from a video clip, the camera relationship can now be established. The establishment is concerned with the spatial relationship on the surveillance area of two adjacent cameras. The relationship can be distinguished into three scenarios: the area covered by the two cameras is continuous, set apart, or partially overlapped.

A time interval parameter is designed in order to establish the camera relationship under the above three scenarios. For two video clips that their direction is not "U", if the difference between the ending time of one video clip and the starting time of the other video clip is smaller or equal than the time interval, the corresponding cameras of two video clips are regarded to be correlated. A camera pair is then established and denoted as (Cam1, Direction1)-(Cam2, Direction2).

The partial results of time intervals two and three seconds are listed in Table 2. The pairs from one to five seconds are counted and sorted as listed in Table 3.

Table 2: Partial results of camera pairs derived with time intervals two and three seconds

Time interval	Camera pairs	
2	(18,L)-(24,L)	(12,L)-(18,L)
	(30,C)-(29,L)	(24,L)-(30,C)
	(12,L)-(18,L)	(15,L)-(16,L)
3	(18,L)-(24,L)	(12,L)-(18,L)
	(06,L)-(12,L)	(12,L)-(18,L)
	(30,C)-(29,L)	(24,L)-(30,C)
	(16,L)-(18,L)	(15,L)-(16,L)
	(24,L)-(29,L)	(24,L)-(30,C)

Table 3: Partial results on camera pairs

Camera Pair	Count
(18,L)-(24,L)	30
(30,C)-(29,L)	28
(12,L)-(18,L)	24
(16,L)-(18,L)	20
(15,L)-(16,L)	20

Next, the previous and next binary relationships of a camera are also counted and sorted according to the camera pairs listed from one to five seconds. Partial results are listed in Table 4.

Table 4: Partial results of camera's previous and next relationships

Camera no.	Previous (count)	Next (count)
06	15 (3)	12 (19)
12	06 (19)	18 (24)
15		16 (20), 06 (3)
16	15 (20)	18 (20)
18	12 (24), 16 (20)	24 (30)
24	18 (30)	29 (16), 30 (16)
29	30 (28), 24 (16)	
30	24 (16)	29 (28)

2.3 Triple camera relationship generation

To strengthen the relationship of two adjacent cameras, a three-camera relationship further be established in this step. If a triple relationship can be established among three cameras, it means that there is a greater probability that the three cameras are adjacent to one another. The generation of triple camera relationship is checking any two camera pairs, denoted as (Cam1, Direction1)-(Cam2, Direction2) and (Cam3, Direction3)-(Cam4, Direction4). The generation is according to the following principle:

1. When Cam2 = Cam3, and Cam2 utilizes the identical video clip to establish the pair of Cam1 and Cam4, Cam2 and Cam4 are found to exist in a triple relationship, which is denoted as (Cam1, Cam2, Cam4).
2. When Cam4 = Cam1, and Cam4 utilizes the identical video clip to establish the pair of Cam3 and Cam2, Cam4 and Cam2 are found to exist in a triple relationship, which is denoted as (Cam3, Cam4, Cam2).

Partial results of generated triple relationships are depicted in Table 5. The triple relationships are used to strengthen the relationship among cameras, no direction is recorded with the triple relationship.

Table 5: Partial results of generated triple camera relationships

Triple camera relationship	Count
(12,18,24)	20
(06,12,18)	17
(16,18,24)	16
(24,30,29)	14

2.4 Relationship combination

By utilizing the above generated triple relationship and binary relationship, the relationships are combined at this step based on the following step.

According to the camera's previous/next relationship established earlier in Table 4, a camera number *P* with the largest value derived from a camera *C*'s previous sequences. *C* is checked with the next sequence of *P*. If *C* is also the largest camera number of *P*, *C* and *P* would be regarded as an adjacent relationship. The initial combination results usually consist of several series of relationship as listed in Table 6.

Table 6: The initial combination results

No.	Relations
1	06, 12, 18, 24
2	15, 16
3	30, 29

The above initial results are further combined via the triple relationship according to the following rule:

1. The last camera number in every serial relationship of the initial results is used to locate the most probable next and the next from the previous/next relationship to form a triple relationship.
2. Being similar to the previous rule that requires locating the previous camera. The first camera number in every series of relationship, and the last probable previous camera from the previous/next relationship, and the camera number further up will form a triple relationship.
3. By focusing on cameras that have not been combined via the above rules. The earlier binary relationship established is used to verify whether the first camera or the last camera of the present combination contain a binary relationship. If so, they are added to the foremost or the last of a series of relationship.

According to the above rules, the final combination results are listed in Table 7.

Table 7: The final combination results

No.	Relations
1	06, 12, 18, 24, 30, 29
2	15, 16, 18

Finally, a comparison is made to any two given adjacent camera numbers to the results of the camera pair (as depicted in Table 3) to ensure the direction of a camera. The direction of a camera *P* with movement "C" is denoted as "`<-, P`"; contrarily, if the direction is "L", it is then denoted as "`P, ->`". The final results are depicted in Table 8.

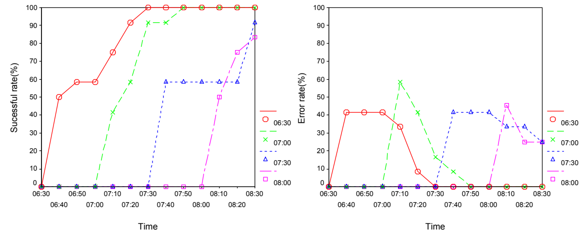
Table 8: The final result

No.	Relations
1	`06,->,12,->,18,->,24,->, <-,30,29->`
2	`15,->,16,->,18,->`

3. Experimental study

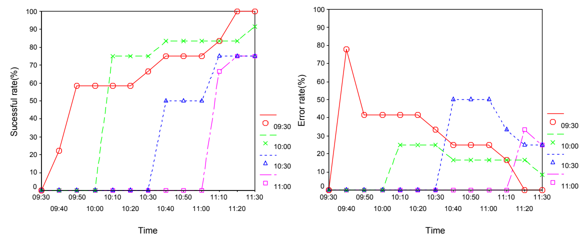
We have conducted an experiment for the automatic establishment of camera relationship on a community surveillance system. There are three roads and a total of 13 cameras installed in the community. The experiment is separated into the peak hours (6:30~8:30am) and the off-peak hours (9:30~11:30am). The relationship is established at the interval of every ten minutes. The correct and error ratios in percentage are used to evaluate the performance of our method.

The experiment results are depicted in Figures 3 and 4. Figure 3 depicts the successful ratio and error ratio. Assume the total number of camera relationship is denoted as *N*, and *M* relationship established in a given time period, in which the correct relationship is denoted as $m1$, and the error relationship is denoted as $m2$, M=($m1$+$m2$). The correct and error ratios equal to $m1/N$ and $m2/N$, respectively.



(a) (b)

Figure 3: Peak hour results



(a) (b)

Figure 4: Off-peak hour results

As can be extrapolated from the experiment charts, regardless of what time period, the volume of a 10-minute video clip will suffice to bring the successful ratio to 50%, and all camera relationship can be identified comprehensively in roughly a matter of 60 minutes; conversely, the error ratio tends to diminish as the timing increases, which is the major advantage of the method the paper presents, whereby through the accumulation of the volume, it becomes easier to weed out the erroneous scenarios. Figure 4 depicts an successful ratio and error ratio graph conducted during the off-peak hours, in which as the volume of moving objects appear less during the off-peak hours, resulting in a relatively lesser camera relationship identified using the method, it takes a longer time of approximately 100 minutes or so in order to locate all correct relationship when compared with that during the peak hours; in addition, when analyzing the first 10 minute of a video clip, only approximately 20% to 30% of successful relationship can be identified.

4. Conclusion

The paper attempts to design an automatic method for establishing the camera relationship by analyzing the video clips of a multi-camera surveillance system. According to the experimental results, the method is able to generate comprehensively camera relationship, and as long as there is a given span of time defined, it is able to detect in a video clip the direction and relationship among the cameras based on the timing and spatial information of a moving object. The proposed method is useful to enhance the practicality and convenience of multiple camera services and applications.

References:

[1] W. M. Hu, X. J. Xiao, D. Xie, and T. N. Tan, "Traffic Accident Prediction Using Vehicle Tracking and Trajectory Analysis," in *Intelligent Transportation Systems*, Vol.1, 2003, pp.220-225.

[2] J. W. Hsieh, S. H. Yu, Y. S. Chen, and W. F. Hu, "Automatic Traffic Surveillance System for Vehicle Tracking and Classification," in *Intelligent Transportation Systems*, Vol.7, June 2006, pp.175-187.

[3] W. M. Hu, M. Hu, X. Zhou, T. N. Tan, J. G. Lou, and S. Maybank, "Principal Axis-Based Correspondence between Multiple Cameras for People Tracking," in *Pattern Analysis and Machine Intelligence*, Vol. 28, April 2006, pp.663 - 671.

[4] J. X. Wang, G. Bebis, and R. Miller, "Robust Video-Based Surveillance by Integrating Target Detection with Tracking," in *Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop,* June 2006, pp.137.

[5] J. Black and T. Ellis, "Multi camera image tracking," in *Image and Vision Computing,* June 2005, pp. 1256-1267.

[6] O. Javed, Z. Rasheed, O. Alatas, and M. Shah, "KNIGHT: A Real Time Surveillance System For Multiple Overlapping And Non-Overlapping Cameras," in *ICME 2003*, 2003, pp.649-652.

[7] O. Javed, and K. Shafique, "Automated Visual Surveillance in Realistic Scenarios," in *IEEE Multimedia*, Vol. 14, Jan.-March 2007, pp.30-39.

[8] G. Jing, D. Rajin and C. E. Siong, "Motion Detection with Adaptive Background And Dynamic Thresholds," in *Information, Communications and Signal Processing,* Dec. 2005, pp. 41 -45.

[9] A. Colombari, A. Fusiello, and V. Murino, "Segmentation and Tracking of Multiple Video Objects", *Pattern Recognition*, Vol. 40, Issue 4, Apr. 2007, pp. 1307-1317.