OBJECTS TRACKING IN A VIDEO SEQUENCE


ABSTRACT

OBJECTS TRACKING IN A VIDEO SEQUENCE. This paper presents the result of implementing a tracking system for identifying objects in a video sequence. The main objective of this research is to keep track of objects movement and their activities which are then analyzed whether the activities related to suspicious activities or not. At this stage the research is concentrated on the keep track of the objects once the objects enter the scene. The objects tracking are done by identify objects’ movement from video sequence using frame by frame analysis. In order to avoid tracking unnecessary objects a method is implemented to eliminate such objects. In this research a method to eliminate such objects is to use spatial objects information. Based on the described method the research shows that objects tracking in a video sequence can be implemented. Moreover, the research is also trying to isolate objects so that the object size and its activities can be analyzed. Finally, this research has been implemented both offline and real-time environment and the results show no significant different.

Keywords: Objects Tracking, Objects Movement, Objects Activities, Video Sequence, Frame.
INTRODUCTION

The challenge for detection and or recognition of objects in an environment for surveillance purposes is really a step ahead. Many groups and or personal researchers have paid serious attention to develop smart and intelligent systems which is very different from the conventional surveillance system. The conventional surveillance system cannot give information into the authority so that it cannot prevent any potential dangerous situation and or crime which might happen. The advance in computer technology particularly in the field of computer vision has made it is possible to develop a smart and intelligent surveillance system. This system can prevent a dangerous situation by feeding information to the authority if exist suspicious objects and or their activities in an area. Many researchers have implemented the idea of smart and intelligent system mostly in specific areas such as parking lot, train stations, crowded movement people, military purposes.

Figure 1. An Intrusion to a Nuclear Power Plant.

One illustration that an active video surveillance system is exceptionally required is in a highly restricted environment. The above figure 1 shows an intrusion to a nuclear power plant in England which happened as a case for the weakness of the conventional surveillance system.

In computer technology, generally there are two (2) methods that are common in computer vision approaches to do the detection i.e. feature based and template matching methods. In feature based method there are several information that can be
are used to detect objects such as shape, color, texture and distinctive features. All of these features are very domain and environment specific that suit for some applications. On the other way, in the template matching method there are two common approaches i.e. fixed and deformable or dynamic template. The fixed template matching techniques are suitable for detecting objects which are not shape deformable or the objects shapes are relatively constant in field of camera view. The deformable or dynamic templates are used for applications where objects are change in shape.

Based on the above classification, this report describes a method of objects tracking based on the feature based. The objects which are tracked are relatively constant in the field of camera view. At this stage, the main feature considered is object shape which is used to filter moving objects by determining their edges and shape. Once the objects are recognized then the objects will be tracked. At the next stage once the object are tracked then the objects’ activities will be recorded and analyzed to determine whether the activities are suspicious or harmful in term of surveillance.

In this report, the methodology described has been implemented for tracking objects in a house and there are no significance differences whether it is implemented as an offline or real time system.

**METHODOLOGY**

The basic system of this research to achieve is called an Active Video Surveillance System. Basically the system is shown in figure 2 below:

![Figure 2. An Active Video Surveillance System.](image-url)
The system captures video data from the environment (surrounding area). This video data is then processed by the system, which in turn, will produce security information i.e. safe or unsafe condition. Safe condition means that there are no any possibly suspicious objects and or activities while unsafe condition means there is a possibility that a suspicious object and or activities might happen.

Video data captured in form of frames is stored into computer memory. These frames are processed in order to find the motion of objects within the video stream. The object detection and tracking techniques have been applied for real-time and offline video stream.

The more detail method of this technique can be described into several stages as shown in figure 3 below:

![Diagram of Object Detection and Tracking](image)

Figure 3. Object Detection and Tracking

At first stage the video data captured from an environment either online or offline is stored into computer memory. There is no special treatment is applied for offline nor online data source. Both the real-time and offline video data captured is split into frames before it is processed. The timer is set up accordingly in order to register the frames.

The second stage in this processing is to register the background image or frame. For the first background i.e. the first image then there is no following activity in term of the frame is not manipulated any further. The reason for this is that there is no any data to be compared with the first image. It is realized that in computer vision
the lighting condition and other environments’ factors influence the accurateness of image capture overtime. In order to reduce the influence of the lighting condition a dynamic background adaptation is carried out. In this research the background adaptation is done by replace the image background with the previous image.

The third stage in this method is to do some enhancement and or data manipulation. The enhancement method applied is taken from some common available methods such as noise removal, contrast and some other popular data manipulation.

The next stage applied is the edges detection method. The method applied here is part of the research i.e. image subtraction method as described in more detail in [13]. Basically image subtraction can be represented as:

\[ \Delta I(i,j) = I_{Cur}(i,j) - I_{Prev}(i,j) \]  

where:

- \( \Delta I(i,j) \) is the image intensity different from 2 consecutive frames.
- \( I_{Cur}(i,j) \) and \( I_{Prev}(i,j) \) represent images intensities for Current and Previous frames respectively.

In this research the method applied is purely by comparing image pixels from two consecutive frames. If there exist a different value between these two frames then consequently it represents motion. The differences between these two frames will be represented as the moving parts in the scene. The destination frame displayed consists of the combination of the moving regions which is then used for edges detection of the objects.

Furthermore, in order to reduce unnecessary edges a threshold is chosen in such a way that image intensity differences \( \Delta I(i,j) \) will not automatically becomes a motion unless the different is greater than the threshold. In this technique, the frames will be scanned twice one from left to right, row after row and secondly from top to bottom, column by column.

After the edges detection process is performed then a process is applied to determine whether an edge belong to an object i.e. boundary of an object. First, the difference between two corresponding pixels (same location \((i,j)\)) of two frames is greater than a threshold. For example: If the \( \text{ABS} [ I_{Cur}(1,11) - I_{Prev}(1,11) ] \) > THRESH) THEN it will be assigned BLACK OR BLUE. Once the edges detection have been applied to the image frames then for successive two frames the images are compared to find the object. In this context, an object is the moving part where edges of the object are different from frame to frame. At this stage the size of the objects is also determined and the objects are assigned the number.

Finally, once the objects have been identified then the objects are tracked based on their size and number. At this stage some objects that are not significant and or cannot be detected in some detail are eliminated. This elimination is done for several
reasons such as there is not enough information that can be derived from small size objects. Furthermore, it is difficult to extract information from small size of objects both activities and or texture data.

IMPLEMENTATION

As it has been described in the previous section the object motion is done using image subtraction method. This method is very straight where once the video data have been captured into images frames then for 2 successive frames the motion is directly processed. In this perspective, the processing is carried out by calculating the intensity differences between previous and current images from the same position as shown by equation (1). The differences show that there exist objects and motion. The algorithm of the objects and motion detection can be described as follows:

% Capture the images
While not end of the frame
    Begin
        % For 2 consecutives images calculate the differences and compare to the threshold
        For X = 0 to MaxWidth
            Begin
                % Calculate the differences
                \[ \Delta I(i,j) = I_{cur}(i,j) - I_{prev}(i,j) \]
                If \[ \Delta I(i,j) < \text{Threshold} \] then
                    \[ I(i,j) = \text{BLACK} \]
                Else
                    \[ I(i,j) = \text{BLUE} \]
                End if
                % Plot the objects and their rectangle into images
                WRITE \[ I(i,j) \]
            End
        End
    End

216
The above algorithm has been applied to several cases, which is one of the cases is in the parking lot. The following figures 4 shows object detection method applied to a working room. In this scenario, the method detects several moving objects and the system tries to isolate the objects by drawing the objects size in rectangular form. Furthermore, the objects are labeled by numbering each object so that a tracking mechanism can be applied.

Figure. 4a, b and c. Several consecutive frames and the objects tracked.
Figure 5 shows the application of the method using some refinement and or enhancement. In this research the same environment is used but in this case the tracking results in a better outcome in term of some number of unnecessary objects are eliminated. By eliminating some unnecessary objects results in less complicated tracking system.

Figure. 5a, b, c and d. Several consecutive frames and the refined objects tracked.
RESULTS AND DISCUSSION

We have described our propose method for an active video surveillance system. This propose system is an effort to relieve security officer from doing a boring job as the system can provide information to security officers about objects in their environment. Furthermore, the information can be classified into safe or unsafe information. In this research, we have shown that the method has the ability to recognize objects and track their moving using computer vision approach. Now, we are in the process of describing the objects identification and classification. Then, we are going to apply some measurement parameters such as physical size, texture and their possibly activities. The described techniques together with skeletonization technique described in [6] will be used to make an active video surveillance system to the real world by exploring the computer vision technique in more comprehensive. At this stage a method of objects tracking in a video sequence has been implemented with promising results as can be seen in figure 4 and figure 5.

FUTURE WORKS

In this research, we have implemented an objects tracking of a video sequence which is part of the requirements required to implement an active video surveillance system. We will explore in more comprehensive way so that the challenge in an active computer vision system can be implemented.

REFERENCES


15. SINGH SAMEER, HADDON JOHN, MARKOU MARKOS, Nearest-neighbour classifiers in natural scene analysis, *Pattern Recognition* 34 (2001)


ELOK AMITAYANI

Sejauh mana kemampuan program anda untuk meramalkan kegiatan objek? Bisakah mengidentifikasi gerakan mencurigakan?

R.B WAHYU

Sesuai dengan parameter yang saat ini sedang saya kerjakan maka program yang saya kembangkan ini harus dapat menganalisa kegiatan objek yang berada pada kawasan yang dimonitor. Hasil akhir dari kegiatan ini adalah smart surveilan system yang mampu mengidentifikasi kegiatan yang mencurigakan seperti jika seseorang bolak balik di area yang dimonitor beberapa kali.

SLAMET

1. Perangkat Lunak apa yang digunakan untuk pengembangan sistem ini?
2. Bila objek yang dicurigai lebih dari satu bagaimana penanganannya?

R.B WAHYU

1. Pada tahap awal percobaan perangkat lunak yang digunakan adalah Matlab. Namun dari hasil percobaan yang dilakukan respon time dari sistem kurang baik (lambat). Untuk mengatasi hal ini sekarang perangkat lunak yang digunakan adalah Microsoft C#

TOPAN

1. Berapa fps(frame per second) optimal yang bisa diperoleh sejauh ini?
2. Jika gambar dikirim melalui fasilitas komunikasi berapa kecepatan dalam kbps yang optimal?
1. Dalam melakukan analisa untuk surveilan jumlah frame tidak terlalu signifikan. Dalam hal ini jumlah fps sangat tergantung pada domain aplikasi seperti jika pengawasan terhadap gerak gerik orang di kawasan terbatas waktu satu atau dua menit monitoring merupakan waktu yang yang cukup untuk mengamati. Untuk sistem yang dikembangkan ini saat ini jumlah frame yang diproses cukup optimal adalah 3 fps.

2. Pada saat ini hasil pengolahan maupun gambar yang direkam belum dikirimkan melalui fasilitas komunikasi. Namun hal ini sangat dimungkinkan karena dengan menggunakan webcam yang sangat sederhana dan dengan resolusi yang tidak terlalu tinggi telah dimungkinkan untuk menganalisa objek yang ada. Kebutuhan akan informasi gambar dan hasil analisa dimungkinkan untuk dikirim melalui fasilitas komunikasi saat ini seperti pengiriman ke hand phone.