CRISP SET IMPLEMENTATION ON VIDEO IMAGES FOR THE APPLICATION OF SURVEILLANCE SYSTEMS

Altahir A. Altahir , Vijanth S. Asirvadam and Patrick Sebastian

Department of Electrical and Electronics Engineering
Universiti Teknologi PETRONAS
Email: vijanth@gmail.com, vijanth_sagayan@petronas.com.my

ABSTRACT

Observing moving objects in far field's video surveillance is one of the main application areas in computer vision. The strong interest in this research direction is driven by creating full automotive surveillance applications. This paper presents implementing a crisp set on video images in order to evaluate human activities in far field's surveillance systems. Reducing the storage capacity in surveillance systems is discussed also in this paper. The concept is based on extracting two powerful attributes from objects motion, namely velocity and pixel frequency distribution. This step followed by combining the measurements mentioned above via crisp set rules in order to evaluate the active section in the image plane and to determine the suitable storing rate. The experimental results proved the efficiency of the novel approach.

Key words – *Set theory, Surveillance System frequency distribution, Image Processing.*

1.0 INTRODUCTION

The implementation of crisp set theory in image processing and pattern recognition appeared in late sixtieth and early seventieth of the last century. The implementation of crisp rules on image at that time was image enhancement, segmentation and edge detection (Altahir *et. al* 2007). The main idea of crisp set mathematics is constructed on two basic components of set theory, the sets themselves and the operations on those sets. Crisp logic defines rules, based on combinations of the sets by these operations.

Due to the requirements of assisting the human operators in the surveillance systems to catch the events of interest and evaluating the suitable storage rate, this paper introduces a novel trend of implementing crisp sets on image processing in order to create semi automotive surveillance system. This concept is based on extracting two powerful measurements from object motion (Altahir *et. al* 2008a, 2008b), then combining the measurements mentioned above via crisp rules in order to evaluate the active section in the image plane. Figure 1 illustrates the basic operations in order to extract the pixel distribution and velocity, and the relation with the fuzzy algorithm. According to the measured values of object velocity and pixel distribution a weight is assigned for each single second in the video stream. Finally based on the dedicated weight a suitable storage rate is allocated to each second in the video data.

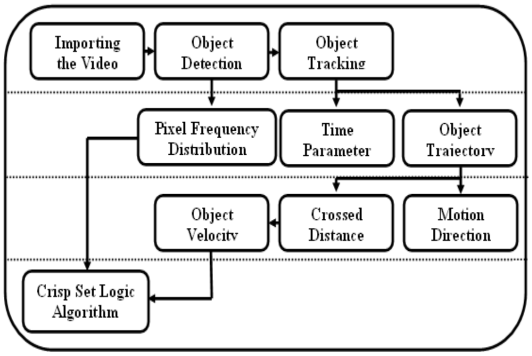

Figure 1: Extracting and combining the pixel distribution and object velocity.

The rest of this paper is organized as follows: section two discusses the object detection. The criteria extraction is given in section three. Section four describes the proposed method for evaluating the activities. Section five describes reducing the storage capacity in online video data. The conclusion appears in section six and it's dedicated to present the graphic user interface for the proposed system.

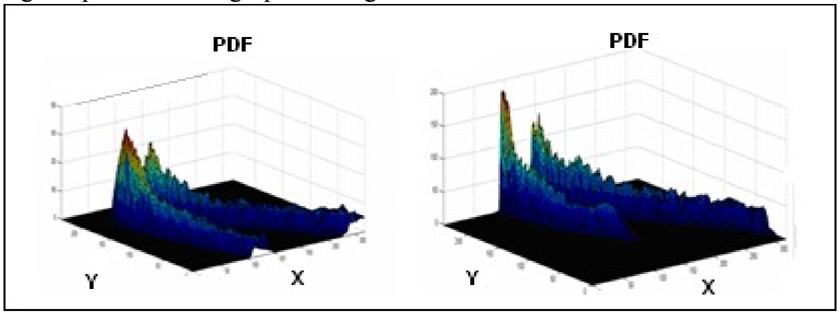

Figure 2: Pixel frequency distribution for the two video samples under study.

2.0 OBJECT DETECTION

Detecting moving object in a video sequence is a crucial part of any video contained analysis applications. In this work the detecting process is required due to the need of extracting the pixels distribution and the object velocity, where the detection process is realized based on static background subtraction method created by (Heikkila and Silven, 1999); to segment each frame into foreground regions {moving objects} and background {the static scene}. The detection process usually followed by thresholding the resulted image with a fixed threshold to evaluate the image pixels. After that the background image is updated, and a closing window is utilized for manipulating the subtracted result.

3.0 CRITERIA EXTRACTION

This section aims to illustrate extracting pixel frequency distribution and objects velocity in order to evaluate human activities. Throughout this work, 530 frames of video sample containing two agents walking, is used for the purpose of evaluating human activities (Altahir *et. al* 2008b),. The evaluation result is presented in the form of visual alarms in the particular region which witnessed the high level of activity. The subsections below dedicated to discuss the extraction process.

3.1 Pixel Frequency Distribution

Pixel frequency distribution is considered one of the most powerful criteria for analyzing the objects motion. So, it used to describe the motion activity with respect to the location. Moreover, it presents the details of the motion activities in a visual understandable manner (Altahir *et. al* 2008a, 2008b).

Equation (1) illustrates the process of calculating the pixel frequency distribution from a mathematic point of view:

$$F(i,j) = \sum_{t=1}^{K} \sum_{i,j=1}^{M,N} I(i,j) \quad (1) \qquad (1)$$

Having F stands for the accumulation of pixel intensities $I(i,j)$ is the pixel value corresponding to the location i, j in a digital image, M and N the number of rows and columns and K is the number of frames.

Figure 2 shows pixel frequency distribution for the video sample under study.

3.2 Velocity

The definition of the term velocity in our work is a measurement describes how fast the object moves from one point in the image plane to another point in the same image plane on pixels bases (Altahir *et. al* 2008a, 2008b), Object velocity is calculated based on the crossed distance and the time needed to achieve crossing the particular distance. Regarding calculating the crossed distance we need first to form the motion trajectory, Figure 3 presents the objects trajectories.

After generating the trajectories, a suitable step between the trajectory points is needed in order to calculate the crossed distance.

The necessity of utilizing time threshold comes from avoiding the ignorable small distance between the successive frames. Moreover, it also considers the details of the motion behavior, for example if we consider finding the crossed distance between the first point of the motion {the very first point in the object trajectory} and the last point, the result is straight line which in most cases differs from the actual crossed distance.

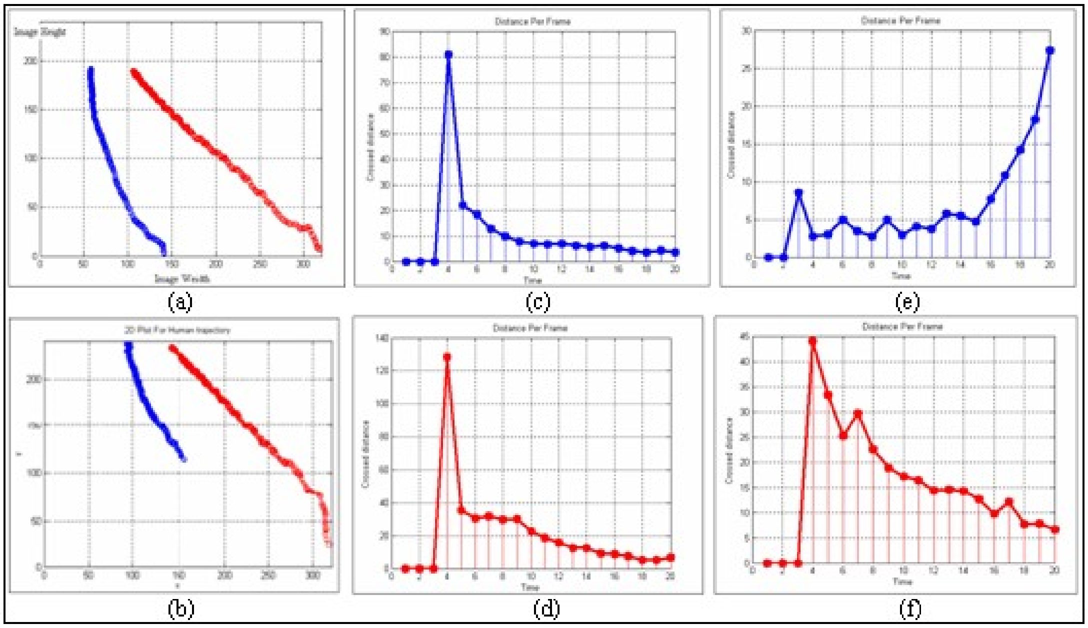

Figure 3: Objects trajectories & the crossed distance for the two for the two video samples. (a) The trajectory for the first video sample. (b) The trajectory for the first video sample. (c) The crossed distance for the first agent in the first video sample. (d) The crossed distance for the second agent in the first video sample. (e) The crossed distance for the first agent in the second video sample. (f) The crossed distance for the second agent in the second video sample.

The mathematical base for measuring the crossed distance is the space between two points in two dimensions space. So crossed distance is calculated according to the Equation (2):

$$D^k = \{d_1, d_2, \ldots d_i \ldots d_n\} \qquad (2)$$

$$d_i = \sqrt{\left(\bar{x}_i - \bar{x}_{i-1}\right)^2 + \left(\bar{y}_i - \bar{y}_{i-1}\right)^2} \qquad (3)$$

Having D is a crossed distances vector, k is the agent index, d_i the distance crossed in frame i, x_i and y_i is the current Centroid coordinates, x_{i-1} and y_{i-1} is the previous Centroid coordinates and n the number of frames.

The results of measured distances for the two agents are presented in Figure 4:

The velocity of the agents is formed based on the calculated distance over the time needed to cross the particular distance. Equation (4) illustrates this concept:

$$V^k = \{v_1, v_2, \ldots v_i \ldots v_n\}$$
$$v_i = d_i / t \qquad (4)$$

Where V^k is the vector containing the velocities, and v_i the velocity in frame i. Figure (5) shows the velocity for the two agents participated in this video sample.

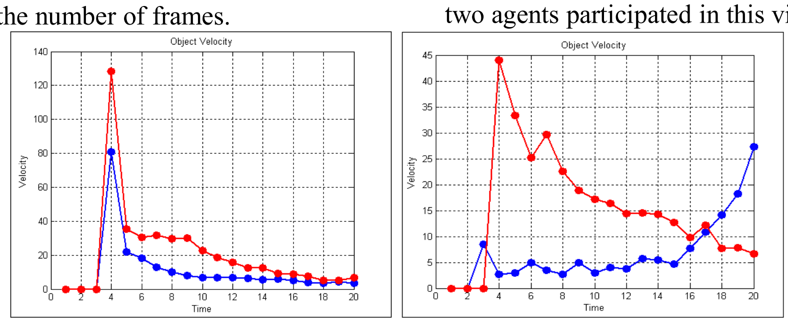

Figure 4: The velocity the two agents

The crisp set algorithm works on the pixel frequency distribution and object velocity. Before passing the inputs to the algorithm, a recalls step is needed. The next section discusses rescaling the inputs and the concept of fuzzy image algorithm in more details.

3.3 Rescaling the selected criteria

The process of evaluating human activities requires rescaling the pixel frequency distribution by choosing the maximum value recorded in each second, moreover it requires also matching up the velocity vectors by means of comparing the current recorded velocities and considering the high velocity during each single second. Figure 5(a) and Figure (b) illustrates the results of rescaling the objects velocities, while Figure 5(c) and Figure (d) shows the results of rescaling pixel frequency distribution.

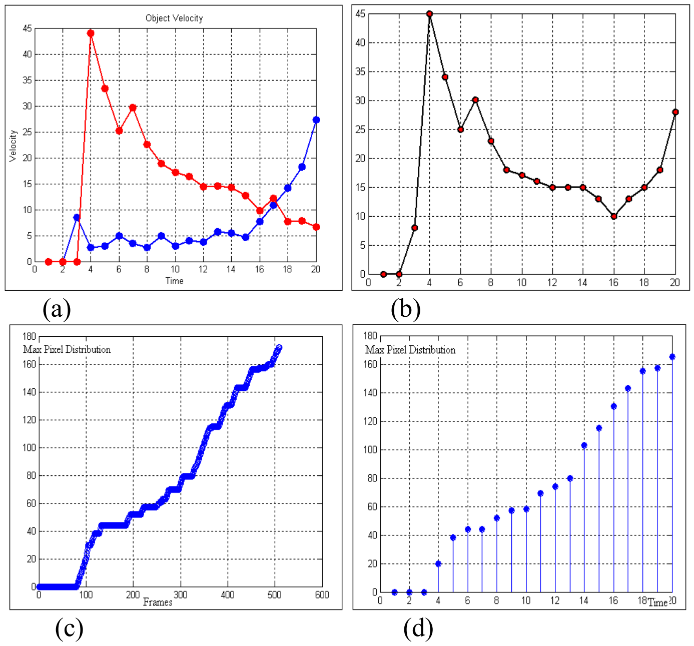

Figure 5: Rescaling the velocity vectors and the pixel frequency distribution

4.0 ACTIVITIES CLASSIFICATION

The crisp set rules algorithm evaluates the rescaled values of the pixel frequency distribution with a normal behavior set N_P, and evaluates the maximum recorded value of velocity with a normal behavior set N_V , in order to determine the highly activities region. Table 1 shows the processes of classifying the active and non active region.

Table 1: *Evaluating the human activities.*

Pixel Frequency Distribution	Velocity	Region Status
0	0	Not Active
$P_i \in N_P$	0	Not Active
$P_i \notin N_P$	0	Active
$P_i \in N_P$	$V_i \in N_V$	Not Active
$P_i \in N_P$	$V_i \notin N_V$	Active
$P_i \notin N_P$	$V_i \notin N_V$	Active

Where, N_P and N_V are the normal behavior set for the pixel frequency distribution and velocity respectively. Constructing the normal behavior sets is a supervised operation relies on the security requirements. The two sets used in this case study are capable of handling any other similar scenarios, while considering the initial parameters like the position of the camera and scene topology.

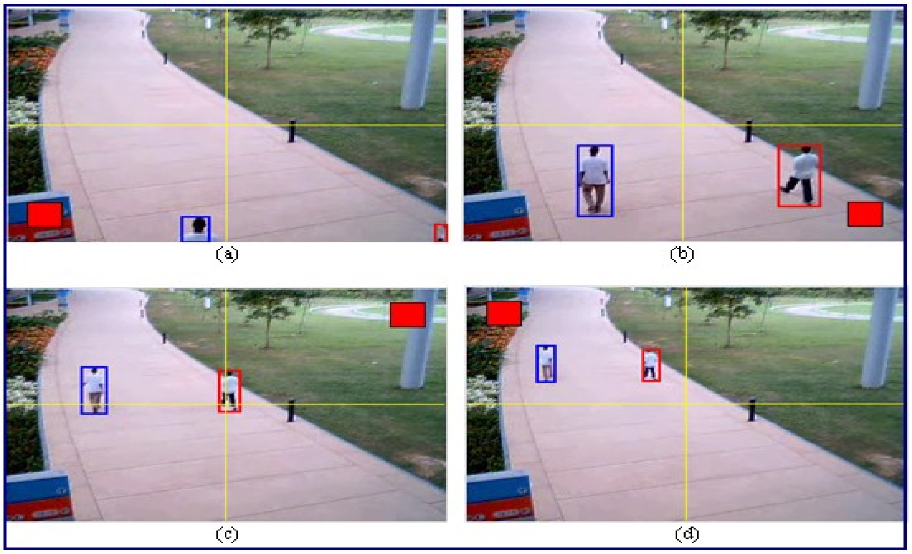

Figure 6: Samples of the visual alarm triggering event.

To trigger the visual alarm, the image plane is segmented into sub zones. The concept of dividing the image plane into group of engrossed sections facilitates the progress of concentrating the effort in a highly sensitive area in the camera view. We implement a based pixel segmenting algorithm to perform this task. Each new coming frame is divided into four zones. After that, we checked the existence of the human in the dedicated zone. Based on the learned activity status, a visual alarm will trigger to assist the security officers to concentrate in the active regions. Figure shows samples of the visual alarm triggering event.

Figure 6(a) presents the entrance of the objects to the scene. The visual alarm appears as a red square and indicates the active region. As explained before the region's activity status is determined based on the velocity and pixel distribution level exhibited in the certain region. Figure 6(b) shows that region two is witnessing high activity more than the rest of the image regions, this fact appeared clearly through the visual alarming system, where it detects that the second agents {red bounding box} presents more activity at that time compared to the first agent {blue bounding box}. Figure6(c) shows that region three is now witnessing high activity more that the rest of the image regions, this fact appeared clearly through the visual alarming system, where still detects that the second agents {red bounding box} presents more activity at that time compare to the first agent {blue bounding box}. Figure 6(d) illustrates the most active region currently is region one and that's because it contains the activities of the two agents participated in this scenario.

5.0 REDUCING THE STORAGE

Reducing the storage capacity is achieved via implementing the set theory logic rules in order to evaluate the suitable storage frame rate for the current events. This process requires rescaling the pixel frequency distribution by choosing the maximum value recorded in each second, moreover it requires also reshaping the velocity vectors by means of considering the high velocity during each single second.

The crisp set algorithm evaluates the values of pixel frequency distribution with two corresponding sets N_P and U_P, and the values of velocity with two sets N_V and U_V in order to determine weights vector W. Shown in Table 2 are the processes of generation the weight vector.

Table 2: *Generating the weight vector.*

W	Pixel Frequency Distribution	Velocity	Storage Rate
0.4	0	0	10
0.6	$P_i \in N_P$	0	15
0.6	$P_i \in U_P$	0	15
0.6	$P_i \in N_P$	$V_i \in N_V$	15
0.6	$P_i \in N_P$	$V_i \in U_V$	15
0.8	$P_i \in N_P$	$V_i \in U_V$	20
1	$P_i \in U_P$	$V_i \in U_V$	25

Where, N_P and U_P stands for normal and abnormal set of values for pixel frequency distribution, N_V , U_V refers to normal and abnormal set of values for object velocity , P_i and V_i the current values for pixel distribution and velocity respectively. Constructing the normal and abnormal sets is a supervised operation relies on the security requirements and differs from user to another. The two sets used in this case study are capable of handling any other scenarios share the scenario under study the initial parameters like the position of the camera and scene topology.

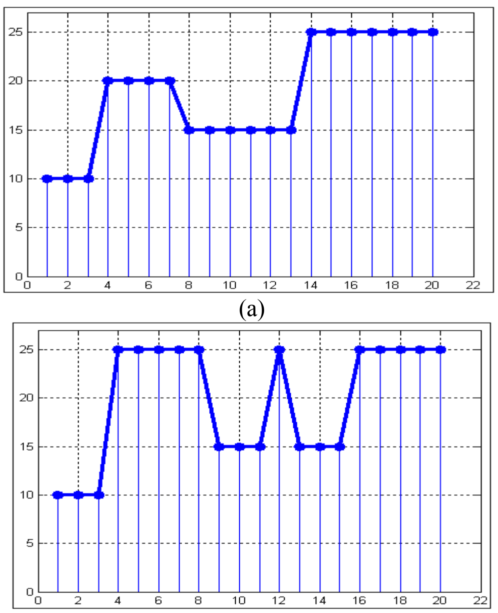

Figure 7: The storage rate per second.

Based on the weight vector we are able to determine a suitable storage rate according to the current events running in the camera view.

Figure 7 shows the suitable storage rate per second. From Figure 7 we are able to conclude by describing the storage rate status over the time where the first three second stored based on 10 frame per second as a storage rate instead of 25 frame per second, from the fourth second tell seventh second stored based on 20 frame per second, from the eighth second tell thirteenth stored based on 15 frame per second and finally the rest of the time stored based on 25 frame per second.

From the discussion above, the proposed approach succeeded in reducing the storage rate based on evaluating the current events and generate a suitable storage rate according to these events without any need to implement any compression algorithms.

6.0 CONCLUSION

This work is motivated by the necessity to create a semi automated surveillance system able to assist the security officers to catch the events of interest from the current scene by observing humans from video stream .The also provides reduction in storage.

The current limitation for this system appears in the tracking part, where the prefect tracking system must avoid occlusion problem which it occurred when one or more object cover another object from the view of the camera, and also must handle complex motion aspects. So the suggested improvement for the current system is implementing a tracking method capable of tracing multi objects in highly active scenes and labeling these objects.

REFERENCES

Altahir. A. Altahir, V. Asirvadam, and P. Sebastian (2007) “Mining Moving Object Attributes Based on Pixel Location for Fixed Camera”. International Conference on Intelligent and Advanced Systems, 2007. pp: 810 – 814.

Altahir, A, Altahir. V, Asirvadam. P. Sebastian and A. Abdalwhab (2008a). “Knowledge Extraction from Human Motion”. International Conference on Computer and Communication Engineering, 13-15, pp 827 – 832.

Altahir, A, Altahir. V, Asirvadam. and P. Sebastian (2008b). “A practical approach for human motion understanding”, The 4th International conference on signal processing & its application (CSPA 08), pp. 145-149.

Heikkila J. and O, Silven (1999). A real-time system for monitoring of cyclists and pedestrians. The 2nd IEEE workshop on visual surveillance. Colorado. pp: 74-81.