CHAPTER FIVE

CRISP SET IMPLEMENTATION ON VIDEO IMAGES

5.1 Introduction

This chapter demonstrates the decision-making process based on the knowledge learned from the video data. The decision is necessary due to the requirements of the alerting system and reducing the storage capacity for the 24 hours surveillance systems. These needs lead us to implement crisp logic algorithm designed for decision-making process. The proposed method derived from evaluating a selective powerful attributes from the set of full frame and inter frame attributes. Two features are chosen to fulfill the requirements; the pixel distribution from the set of full frame attributes and the velocity from the set of inter frame based attributes. The ambition from generating decisions is to help human operator in catching the events of interest and reducing the storage capacity in 24 hours surveillance systems. It’s important to mention that the final decision must be taken by the human operator who controls the validity of the proposed decision. Reducing the storage capacity is achieved via implementing the crisp logic rules in order to evaluate the suitable storage frame rate for the current events.

The two goals mentioned in the previous section require rescaling the pixel frequency distribution by choosing the maximum value recorded in each second, moreover it require also reshaping the velocity vectors by means of considering the high velocity during each single second.

The next sections are dedicated to illustrate the experimental results gained by relaying on two motion scenarios, the first scenario is single human and consists of three different motion aspects. The second scenario is two human and it also consists of three different motion aspects. Section 5.2 is devoted to illustrate the rescaling step for the selective criteria for the purposes of alerting system and reducing the storage capacity, section 5.3

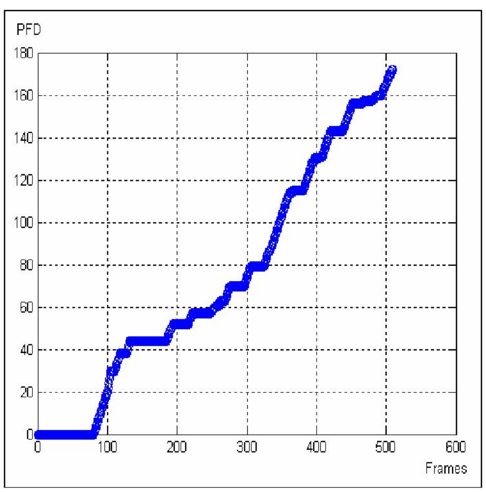
discusses the alerting system unit, while section 5.4 is devoted to illustrate the reducing storage capacity approach based on the current events exhibited by the objects in the scene.

5.2 Rescaling the Selective Criteria

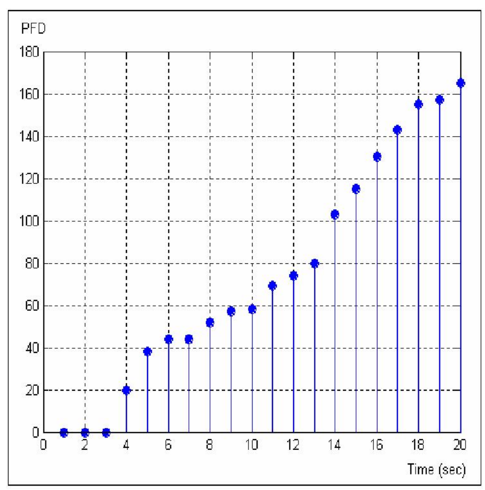
The process of evaluating human activities and reducing the storage capacity requires rescaling the pixel frequency distribution and the velocity vectors in order to prepare the information carried in those two values type for the evaluation process. The next two sub sections illustrate this concept in more details.

5.2.1 Rescaling the pixel frequency distribution

Rescaling the pixel frequency distribution achieved based on selecting the maximum value recorded in each second, or choosing the maximum recorded value during each 25 successive frame. An example of rescaling the pixel frequency distribution is presented in Figure 5.1:



(i)



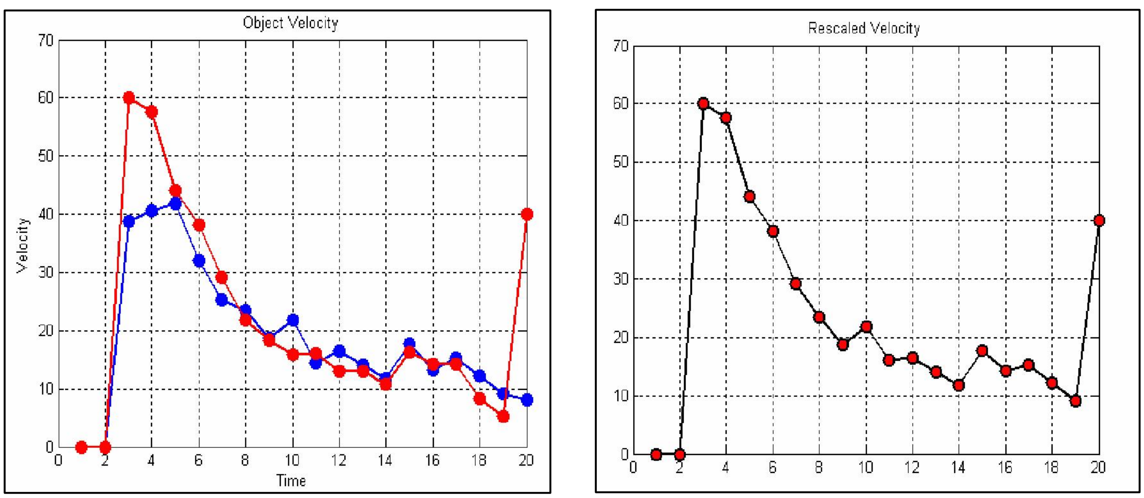
(ii)

Figure 5.1: Rescaling the pixel frequency distribution. (i) The maximum values for the pixel frequency distribution per image frame. (ii) The maximum values for the pixel frequency distribution per second.

Figure 5.1(i) shows the measured values for the maximum pixel frequency distribution over the frame sequence, while Figure 5.1(ii) shows the measured values for the maximum pixel frequency distribution per second bases. Hence, for each second there is a unique pixel distribution value prepared to be an input for the evaluation processes.

5.2.2 Rescaling the velocity vector

The process of evaluating human activities and reducing the storage capacity requires also matching up the velocity vectors by means of comparing the current measured velocities and considering the high velocity during each single second. In addition to comparing the measured velocities, the activity classification needs to identify the current location of the object who registered the matched velocity, with respect of the pre defined image segments. Figure 5.2 shows an example of matching the velocity vectors for two agents walking.



(i) (ii)

Figure 5.2: Rescaling the velocity vectors. (i) The velocity for the two agents participated in the video sample per second. (ii) The result of comparing the velocities for the two agents.

Figure 5.2(i) shows the velocity for the two agents participated in the video sample over the time based on inter framing concepts, while Figure 5.2(ii) shows the result of comparing the velocities for the two agents and it displays the maximum velocity. Hence, for each second there is also a unique velocity value prepared to be an input for the evaluation processes.

5.3 Activity Classification

The crisp 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 5.1 shows the processes of classifying the active and non active region.

Table 5.1: Evaluating the human activities.

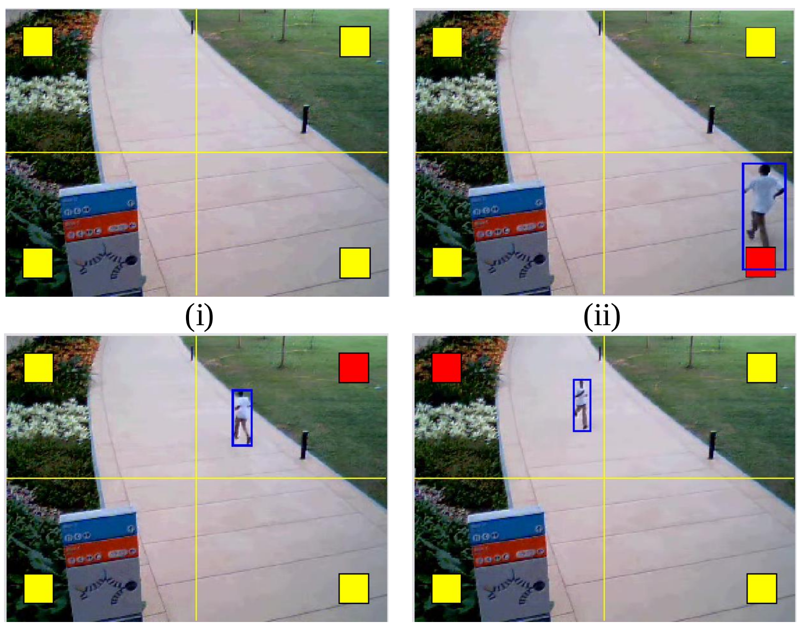
Pixel Frequency Distribution	**Velocity**	**Active region**
0	**0**	**Not Active**
$P_i \in N_P$	**0**	**Not Active**
$P_i \notin N_P$	**0**	**Active**
$P_i \in N_P$	$V_i \in N_V$	**Not Active**
$P_i \in N_P$	$V_i \notin N_V$	**Active**
$P_i \notin N_P$	$V_i \notin N_V$	**Active**

Where, Pi is the re scaled values of Pixel frequency distribution; Vi is the re scaled values of the velocity vector; 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.

Regarding triggering 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 implemented based pixel segmenting algorithm to perform this task. Each new incoming frame is divided into four zones. The segmenting process is followed by determining the location of the object of interest with respect to the image zones. Finally, based on the learned activity status, a visual alarm will trigger to assist the security officers to concentrate in the most active regions. The next sections illustrate the gained results from different case studies.

5.3.1 Activity Classification Based on Single Human Case Study

As stated earlier, the single human case study consists of three video samples describe different motion aspects. For the purpose of classifying the human activity with respect to the image segments, each frame is divided into four sub regions. Figure 5.3 shows the visual alarm while considering a single human running.



(i) (ii) (iii) (vi)

Figure 5.3: Triggering the visual alarm single human case study. (i) No alarm. (ii) The alarm indicates that the active segment is zone 4. (iii) The alarm indicates that the active segment is zone 2. (vi) The alarm indicates that the active segment is zone 1.

Having, the yellow square refers to not active status and the red square refers to active status. Figure 5.3 (i) shows that no activity at all and all the image segments is ideal.

Figure 5.3 (ii) shows the entrance of the object of interest from region four sides, where the alarms changed to active (red) in this segment. Figure 5.3 (iii) shows the object of interest in zone two due to his motion direction, where the alarms changed to active (red) in this segment, and zone four alarm return to ideal (yellow). Figure 5.3 (vi) shows the object of interest in zone one, where the alarms changed to active (red) in this segment, zone two and four alarms return to ideal (yellow).

5.3.2 Activity Classification Based on Two Human Case Studies

Two human, case study consists of three video samples describe different motion aspects. For the purpose of classifying the human activity with respect to the image segments, each frame is divided into four sub regions. Figure 5.4 shows the visual alarm while considering two human walking.

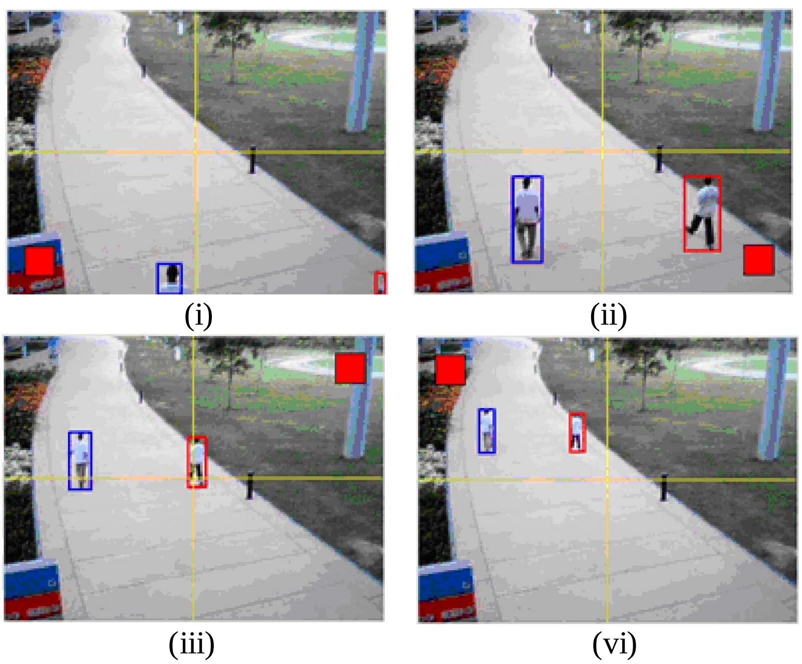


Figure 5.4: Triggering the visual alarm {two human} case study. (i) No alarm. (ii) The alarm indicates that the active segment is zone 4. (iii) The alarm indicates that the active segment is zone 2. (vi) The alarm indicates that the active segment is zone 1.

Not like the previous case, this example shows the different method for expressing the region activity, where the red square indicates the most active region based on the current events, while the ideal and normal regions left without any visual signs.

Figure 5.4(i) present the entrance of the object of interest and the visual alarm appeared as a red square indicates the active region. As explained before the region activity status is determined based on the velocity and pixel distribution level exhibited in the certain region. Figure 5.4(ii) shows that region four is witnessing high activity more that the rest of the image regions, this fact appeared clearly through the visual alarming system, where it detects that the second agents, presents more activity at that time compare to the first agent. Figure 5.4(iii) shows that region two is now witnessing high activity more that the rest of the image regions, this fact appeared clearly through the visual alarming system, where still it detects that the second agents, red bounding box presents more activity at that time compare to the first agent {blue bounding box}. Figure 5.4(vi) 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.

Classifying the activities 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. The simulation scenarios used in this case study allows us to evaluate different human motion aspects. Furthermore, these results show the ease of observing the human activities especially in the case of far field video surveillance.

5.4 Reducing the Storage Capacity

Reducing the storage capacity is achieved via implementing the crisp logic rules in order to generate multi storage rates according to the current activity exhibited by the objects of interest. As mentioned in the previous section, the process of evaluating the current activity 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 output from rescaling process is fed to the crisp logic algorithm. The crisp algorithm works on second bases to evaluate the values of pixel frequency distribution with two corresponding sets, normal pixel distribution set and abnormal pixel distribution set and the values of velocity with two sets, normal velocity set and abnormal velocity set in order to determine weights vector W. Based on the weights vector values generated from the crisp algorithm, a suitable storage rate is determined to match the current activity exhibited by the objects in the scene. Table 5.2 presents the details of the generation the weight vector.

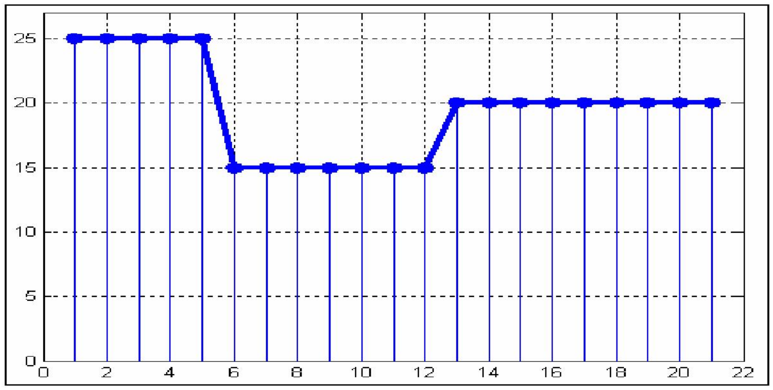
Table 5.2: Generating the weight vector.

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	$P_i \in U_P$	$V_i \in N_V$	**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**

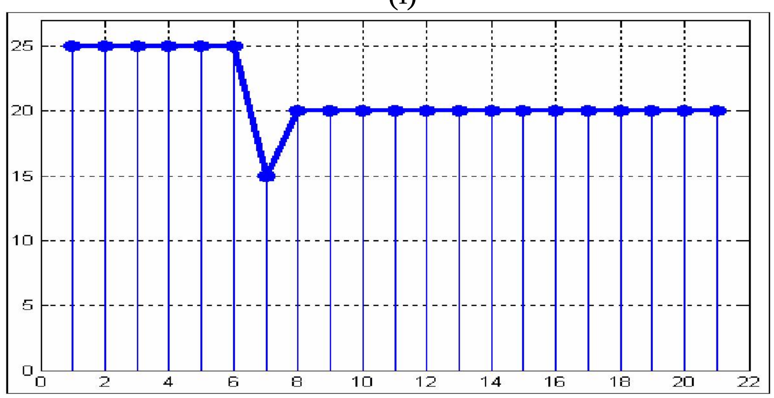
Where, N_P and U_P stands for normal and abnormal set of values for pixel frequency distribution, N_V , U_V refers to normal and abnormal set of values for object velocity , P_i and V_i the current values for pixel distribution and velocity respectively. Constructing the normal and abnormal sets is a supervised operation relies on the security requirements and it may differ from user to another. The next sub sections illustrate the gained results from different case studies.

5.4.1 Reducing the Storage Capacity Based on Single Human Case Study

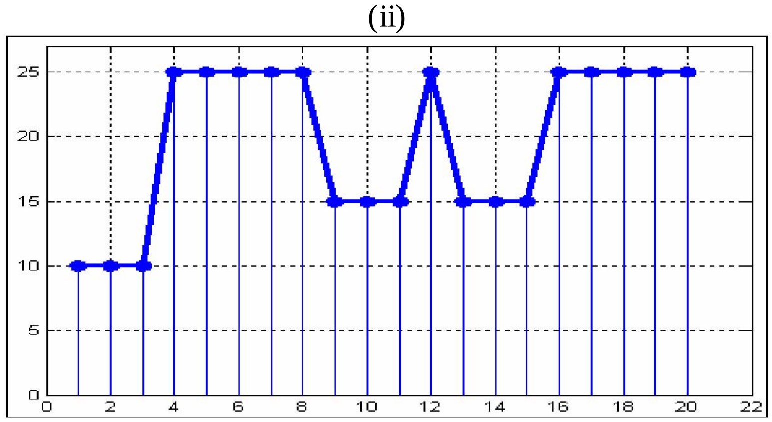
Three different video sequences are used in this case study to describe different single human motion aspects; each one of these video samples contains 530 frames and the total time for each video sample is 20 second. The promise in this section is to calculate the weights vector which allows determining suitable storage rate according to the current events. The results of applying the weights vector produced based on Table 5.2 for the three video samples appeared in Figure 5.5.



(i)



(ii)



(iii)

Figure 5.5: The storage rate per second for the single human case study. (i) The suitable storage rate for the first video sample. (ii) The suitable storage rate for the second video sample. (iii) The suitable storage rate for the third video sample.

From Figure 5.5 concludes by describing the storage rate status over the time for the three video samples. For the first video sample, the first five second stored based on 25 frames per second as a storage rate which it equal to the capturing device frame rate, from the sixth second till twelfth second stored based on 15 frame per second instead of 25, from the thirteenth second till the twenty-one are stored based on 20 frame per second. For the second video sample, the first six seconds stored based on 25 frames per second as a storage rate which it equal to the capturing device frame rate, the second number seven stored at 15 frames per second, the rest of the video stored at 20 frames per second instead of 25. For the third video sample, the first three second stored based on 10 frames per second as a storage rate, from the fourth second tell second number eight eighth stored based on 25 frame per second, from the second number nine tell twelve stored at 15 frame per second, followed by one second stored at 25 frame per second, the second number thirteen up to fifteen stored at 15 frame per second, the rest of the video stored at 25 frames per. Table 5.3 provides a detailed description for the storage rate for the three video samples.

Table 5.3: The storage rate approximations for single human case study.

Video sample	Number of Frames	Length / Second	Required memory space	Required memory space after reduction	Reduction %	Expected rate / minute
1	**525**	**21**	**1050 MB**	**820 MB**	**21.9 %**	**2343 instead of 3000 MB**
2	**525**	**21**	**1050 MB**	**890 MB**	**15.2 %**	**2543 instead of 3000 MB**
3	**500**	**20**	**1000 MB**	**790 MB**	**24.8 %**	**2370 instead of 3000 MB**

5.4.2 Reducing the Storage Capacity Based on Two Human Case Study

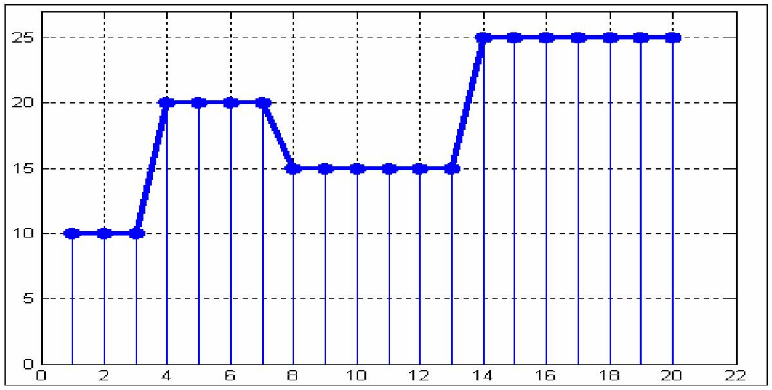
Three different video sequences are used in this case study to describe different two human motion aspects; each one of these video samples contains 500 frames. The total time for each video sample is 20 second [Altahir A. Altahir et al, 2008a].

The motivation here is to calculate the weights vector which allows determining suitable storage rate according to the current events. Figure 5.6 shows the results of applying the weights vector produced based on Table 5.2 for the three video samples.

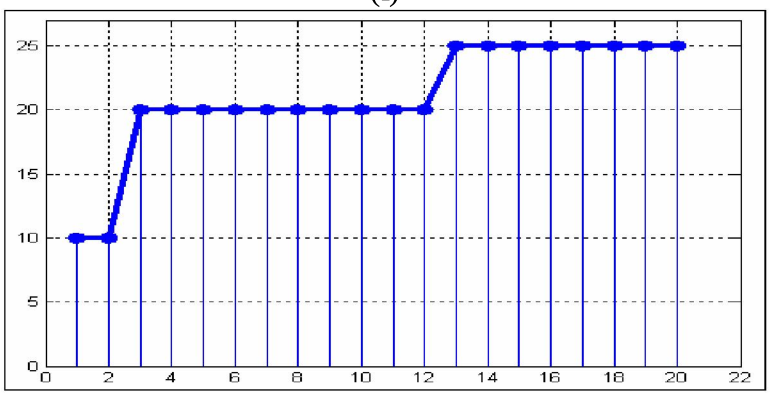
Figure 5.6 describe the storage rate status over the time for the three video samples. For the first video case study (Figure 5.6(i)), the first three second stored at rate of 10 frames per second instead of 25 frames per second, from the second number four up to seven stored at rate 20 frames per second, second number eight up to thirteen stored at 15 frame per second, the rest up to twenty stored at 25 frames per second.

For the second video sample(Figure 5.6(ii)),, the first two seconds stored based on 10 frames per second as a storage rate, the next ten seconds stored at 20 frames per second, the rest up to twenty stored at 25 frames per second.

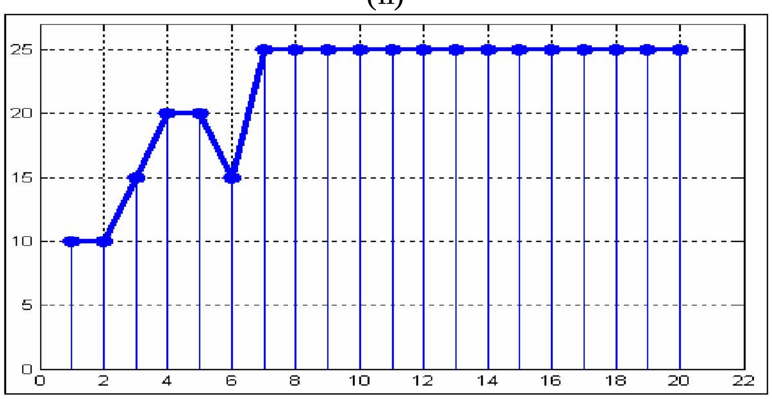
Finally the third video case study (Figure 5.6(iii)),, the first two seconds stored based on 10 frames per second as a storage rate, second number three stored at 15 frames per second, second number four and five stored at 20 frames per second, second number six stored at 15 frames per second, the rest up to twenty stored at 25 frames per second.



(i)



(ii)



(iii)

Figure 5.6: The storage rate per second for the {two human} case study. (i) The suitable storage rate for the first video sample. (ii) The suitable storage rate for the second video sample. (iii) The suitable storage rate for the third video sample.

Table 5.4: The storage rate approximations for two human case study.

Video sample	Number of Frames	Length / Second	Required memory space	Required memory space after reduction	Reduction %	Expected rate / minute
1	**500**	**20**	**1000 MB**	**750 MB**	**25 %**	**2250 instead of 3000 MB**
2	**500**	**20**	**1000 MB**	**840 MB**	**16 %**	**2520 instead of 3000 MB**
3	**500**	**20**	**1000 MB**	**880 MB**	**12 %**	**2640 instead of 3000 MB**

Table 5.4 provides a detailed description for the storage rate for the three video samples. From the discussion above, it can be seen that the proposed approach show that it is possible to reduce the storage capacity by considerable amount of memory space for the different video samples used in section.

5.5 Summary

This chapter is considered as practical implementation of crisp image processing in two directions; the first one is assisting the human operators in the traditional surveillance system to catch the events of interest via visual alarming unit, while the second direction is reducing the storage capacity in 24 hours surveillance systems. The concept is based on extracting selective attributes from object motion, followed by combining these attributes via crisp image rules. The results gained from the two directions are considered strong enough to prove the validity of the proposed method.