Crowd Behavior Analysis under Cameras Network Fusion using Probabilistic Methods

P. Drews Jr, J. Quintas, J. Dias
Institute of Systems and Robotics
University of Coimbra
Coimbra, Portugal
pdrews, jquintas, jorge@isr.uc.pt

M. Andersson, J. Nygårds, J. Rydell
FOI
Swedish Defence Research Agency
Linköping,Sweden
maria.andersson, jonas.nygards, joakim.rydell@foi.se

Abstract – *The use of cameras in surveillance is increasing in the last years due to the low cost of the sensor and the requirement by surveillance in public places. However, the manual analysis of this data is impracticable. Thus, automatic and robust methods to processing this high quantity of data are required. This paper proposes a framework to address this problem. The crowd analysis is achieved in camera networks information by using the optical flow. The Hidden Markov models and Bayesian Networks are compared to understand the agents behavior in the scene. The experimental results are obtained for several sequences where fight and robbery occurs. Results are promise in order to get an automatic system to find abnormal events.*

Keywords: Behavior Analysis, Sensor Fusion, Hidden Markov Models, Bayesian Networks.

1 Introduction

The typical large area surveillance system is characterized by a large network of CCTV cameras, all connected to a control room, where a human operator performs the difficult task of monitoring them all. This fact leads to a situation far from the desired one, since the operator can only pay attention to a little fraction of what is showed on the monitors. Consequently, the task of detecting abnormal behavior in these environments is impossible to be performed in short time, forcing surveillance video to be often used for post-analysis when these situations occur.

Commonly, surveillance systems are installed in public spaces, covering large areas, where a great number of people populates camera's fields of view. Thus, the system's operator sees his job getting more difficult for identifying an abnormal behavior but also it increases the interest for crowd behavior analysis.

Crowd behavior analysis has been target for different studies in a diversity of areas, since social studies until simulation and graphic generation fields. All the involved fields contributes with models and approaches to characterize crowd behavior. In computer vision, this topic is being considered as a new area of interest by the research community, given its potential for creating new application domains, such as automatic detection of riots or chaotic acts in crowds and location of abnormal regions in scenes. A more extensive consideration about related works is presented in the next subsection.

In this article, we have studied two scenarios with a camera network in each one: a robbery at an automated teller machine (ATM) area and fights in an open space (which is assumed to be at an airport scenario). We propose a framework to process this data in order to extract observations, and feed these observations to a Hidden Markov Model (HMM) and a Bayesian Network (BN). Because of the lack of training data, the learning stage for both methods relies on knowledge and experience of what often is associated with normal behavior for the respective scenario. The results for each method are obtained from the likelihood of normal behavior, traducing the probability of an observation being normal given a set of input features.

As a consequence, this paper makes a comparison study between the two methods when performing the task of classifying the observations into normal or abnormal behavior. In Section 2, the description of crowd analysis is briefly presented. Section 3 treats the crowd behavior modeling using both methods (HMM and BN). In Section 4, we present the experimental results which led us to the comparative between the methods. Conclusions and future work are discussed in sections 5 and 6. The last section of this article, Section 7, was reserved for acknowledgments.

1.1 Related works

Crowd behavior analysis has been an important topic of research in the last years. In computer vision community, there are two main approaches to solve this problem. The object-based approach considering the crowd as a collection of individuals [20, 15]. It requires perform segmentation or detection in each object to understand the crowd behavior. The work proposed by [13] detects persons in the scene considering difficulty such as occlusion. The work proposed by [8] finds the movement of each individual person in the crowd using a Bayesian framework. Although the efforts to develop the object-based techniques, there are a lot of intrinsic difficulty related to this kind of technique. They are related to occlusion, segmentation and tracking in images.

The approaches that considering the crowd as a whole have gained importance in the last years [5, 2]. It considers the crowd as a global entity and analysis it in medium to high density scenes. In the work proposed by [17], they use scene modeling technique to capture features to crowd behavior analysis, instead of tracking individual objects. Recently, a method based in Social Force Model [10] was proposed to analyze crowd behavior [14]. It uses a set of particle to estimates the interaction between individuals. [3] proposed a hybrid approach from tracking humans in very dense crowds. In [19] an important review of the crowd analysis is made, it shows the recent research trends and approaches from different research communities.

In previous work [4], optical flow is used in images and counting people in thermal data are proposed as a feature to understand the crowd behavior. Moreover, the HMM is used to obtain the behavior information in the scene. HMM [16] as a tool to modeling behavior is used in the work of [7]. In this work, facial action behavior is modeled. In [5], the crowd behavior is modeled using HMM. On the other hand, the Bayesian Network is a probabilistic method used to understand behaviors. In the work of [9], a framework to understand scenes using multi camera is proposed, it uses Bayesian Networks for behavior analysis based in events.

2 Crowd Analysis

In this paper, two features are acquired in order to detect crowd behavior. The crowd size and crowd activity are estimated using film sequences acquired by a set of cameras. In the next sections these two methods are detailed.

2.1 Estimation of Crowd Size

By detecting and tracking people an estimate of the crowd could be based on a simple count. However in really crowded situations detecting individual persons is difficult. For this reason a more basic measure of crowd size is chosen. The crowd size observations are based on the area of detected foreground in the control area. The size of the control area is chosen so that the area of the detected foreground pixels reasonably approximates the number of people in the control zone given a constant calibration factor (*i.e.* the typical projected area of a person on this distance). For the background and foreground estimation the background is estimated by an approximate median estimate as in [12, 6]. The background update is further masked with a weighting mask based on the detected foreground to avoid static people queuing to blend in with the background. The foreground-probability is then detected as $p_f = 1 - p_b$, where p_b is based on a Gaussian model with the background image as mean and a covariance estimated either as a quartile approximation [6] or as a regular covariance.

Figure 1 presents the counting of people within an area close to one ATM machine, where one robbery takes place. The y-axis represents number of people and the x-axis represents time. In the beginning there is one person close to the ATM and after there are two persons close to the ATM. After the robbery, the robbed person is chasing the thief, and they disappear from the camera view. This information together with movements information in the whole area (not only close to the ATM) can give important information on the event. More details about this scenario is presented in the Section 4.

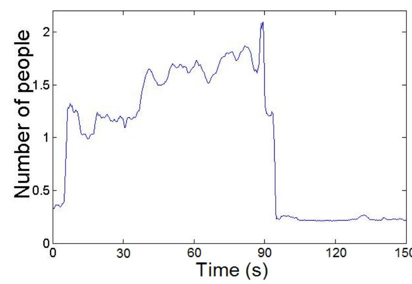


Figure 1: Estimation of crowd size at the ATM scenario.

We also use the counting of people to get information on changes of the crowd size in the short time perspective. By comparing the crowd size from one time to another a measurement of the degree of movements in the crowd can be obtained. A large change may imply that several people are entering or leaving the crowd. It may also imply that there are many movements in the scene and occlusion and shadows contribute to the change in crowd size. This information can be used to complement the activities information from the optical flow (see below). In the work of Andersson *et. al.* [4], the crowd size is estimated using TIR (Thermal InfraRed). This method is used to estimate the crowd size in fight scenarios.

2.2 Estimation of Crowd Activity

The level of activity in a scene, *i.e.* the extent to which persons walk, run, wave their arms etc., can be coarsely estimated by measuring the optical flow in one or more views of the scene. The optical flow in a specific view is measured by computing the apparent motion of each pixel from one video frame to the next. This results in a vector field, where the length of each vector corresponds to estimated magnitude of motion at a certain position in the image. The activity measurement is obtained as the sum of squared motion magnitudes, either in a region of interest or over the entire image. The reason for squaring the magnitudes is that very quick movements which only cover a small part of the image (*e.g.* movements which may occur in a fight) should affect the estimate to a relatively large degree. If the magnitudes are used directly, such small movements are typically drowned among the large number of smaller-magnitude vectors arising from *e.g.* walking persons.

When measuring the activity level in one single view, movements from or toward the camera does not affect the estimate to a very large extent. Hence it is preferable to use multiple cameras. Since the optical flow is calculated by comparing pairs of images, the approach obviously requires that the cameras are stationary.

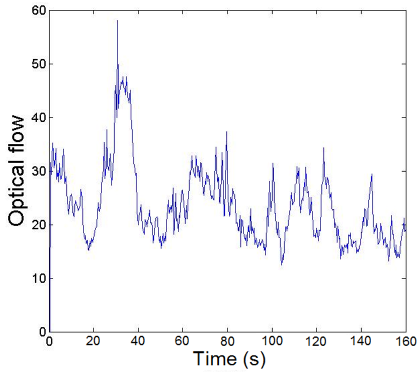


Figure 2: Optical flow from one of the cameras for the ATM scenario.

3 Crowd Behavior Modelling

In this section, we describe two methods used to discover the crowd behavior. They are based in probabilistic approaches. The Bayesian theory gives us the possibility to deal with incomplete data and uncertainly. It makes predictions on future events and provides an embedded scheme for learning.

Specialized models are included in the Bayesian framework, they are known under names as Hidden Markov Models (HMMs), Kalman Filters, Particle Filters and, more generically, Bayesian Nets. This models have been used in a broad range of technical applications. Recent findings indicate that Bayesian models can be useful in modeling of cognitive processes [18]. Research on the human brain and in its computations for perceptions that shows Bayesian methods have proven successful in building computational theories for perception and sensorimotor control [11].

3.1 Hidden Markov Model

In this approach, we propose to process the data from a set of distributed optical sensors (visual and thermal infrared) in order to extract binary observations describing the crowd and feed these observations to a discrete HMM, where the hidden states represent the behavior of the crowd. The approach has been investigated also for other scenarios; see [4].

The HMM is a doubly embedded stochastic process which has an underlying stochastic process that is not observable. The underlying process can be observed through another stochastic process that produces sequences of observations [16]. The states represent some unobservable conditions of the system. In each state there is a certain probability of producing any observable system outputs together with a probability indicating the likely next states. The HMM (λ) is described by the following parameters:

$$\lambda = (A, B, \pi, S, O), \tag{1}$$

where A is the probability distribution of state transitions, B is probability distribution of observations in each state, π is the initial state distribution, S are states in the model and O are discrete observation symbols per state. The parameters of A, B, and π are obtained by training λ on relevant training data.

We use λ to model the normal behavior of a crowd. A high likelihood for a certain observation sequence O_S indicates that the crowd behavior is likely to be normal. A low likelihood indicates that the behavior is abnormal. Observation symbols that represent a crowd in this study include crowd size and movements among people in the crowd. The observations come from distributed optical sensors (visual and thermal infrared cameras).

3.2 Crowd Behavior Analysis using Hidden Markov Model

Detecting and tracking people in crowds is a challenging problem because of occlusions and difficulties in segmenting individuals properly. The idea with the HMM approach is uses observations from the crowd/scene which are not given detailed positions. We do not identify specific persons, or determine their exact positions. We will obtain quite rough decisions on the crowd. These rough decisions can serve as alerts to security operators, who can take a closer look on the specific video and decide if something needs to be done.

Sensor data always include uncertainties. It is important to reduce the uncertainties as much as possible when deriving the observations at the sensor. At the high-level fusion process (HMM) it is possible to further reduce the uncertainties by the combination of observations from the different sensors that have different possibilities to observe the event.

Normal crowd behavior often corresponds to relatively calm movements associated with walking and standing. There should seldom be persons running or waving strongly with arms and legs. In special cases the crowd should not be dense and/or large. The observations are binary and extracted from sensor data using the methods described in section 2. Table 1 presents the observations that we use.

Table 1: Crowd observations.

Observation	Explanation of the observation
O1	Normal activities
O2	Increased activities
O3	Strongly intense activities by many
O4	Small crowd or no crowd
O5	Large crowd
O6	No fast changes in crowd size
O7	Fast changes in crowd size

O1, O2 and O3 are obtained by calculating the optical flows in the visual cameras. O4 and O5 are obtained by calculating the number of people in the scene. In the airport scenario (outdoor scenario) this is done with data from the thermal infrared camera. O6 and O7 are used only in the airport scenario and are based on the thermal infrared data. In the ATM scenario the crowd size is calculated with data from the visual camera, since we had no thermal infrared camera for that case. It is of course advantageous to have also thermal infrared cameras since they can give accurate observations also in poor light conditions.

Since we do not have enough recorded training data, we have derived training data based on knowledge and experience of what often is associated with normal behavior the airport and the ATM. The training is performed by using the expectation-maximization (EM) algorithm.

We have used an ergodic model with two states S1 and S2. S1 refers to calm motions (standing and walking), and S2 refers to slightly increased activities (predominantly walking), still belonging to normal behavior. Table 2 and Table 3 present the HMM parameters that were obtained from the training for the fight (F) and robbery (R) respectively.

Table 2: Initial state probability distribution (π_i) and transition probability distribution (a_ij) for the two cases.

Case	π_{S1}	π_{S2}	a_{11}	a_{12}	a_{21}	a_{22}
F	0.98	0.02	0.45	0.55	0.26	0.74
R	0.86	0.14	0.54	0.46	0.97	0.03

Table 3: Observation probability distribution, B, for the two cases.

B	O1	O2	O3	O4	O5	O6	O7
$B_{S1,F}$	0.66	0.06	0.01	0.11	~ 0.00	0.16	~ 0.00
$B_{S2,F}$	0.24	0.08	~ 0.00	0.33	0.01	0.34	0.01
$B_{S1,R}$	0.81	0.07	~ 0.00	0.10	0.01	-	-
$B_{S2,R}$	0.18	0.26	~ 0.00	0.49	0.06	-	-

Equation 2 is used to calculate the likelihood of normal behavior, where α_t represents the Forward algorithm [16]:

$$\log[P(O|\lambda)] = -\sum_{t=1}^{T} \log \frac{1}{\sum_{i=1}^{N} \alpha_t(i)}. \qquad (2)$$

3.3 Bayesian Network Model

Bayesian networks are applied in cases where there are uncertainty in the data. It is also used when we know certain conditional probabilities and are looking for unknown probabilities. Formally, it is a probabilistic model that represents a set of random variables and their conditional independences via a directed acyclic graph. Edges in this graph represent conditional dependencies and nodes which are not connected represent variables which are conditionally independent of each other. Each node is associated with a probability function that takes as input a set of values for the node's parent variables and gives the probability of the variable represented by the node. Bayesian networks offer the possibility to represent dependencies, parameters and their values intuitively understandable.

3.4 Bayesian Network for Crowd Behavior Analysis

We proposed a Bayesian network shown in the Figure 3 to modelling crowd behaviors. This graphic model could be represented by the Equation 3. It shows the dependencies as a joint distribution and its decomposition while omitting the conjunction symbol $\wedge$.

$$P(B\ E\ In) = P(B)P(E|B)P(E|In). \qquad (3)$$

The *input* variables are the set of nodes shown in the Figure 3 that represents the data obtained from crowd analysis, as described in the Section 2.

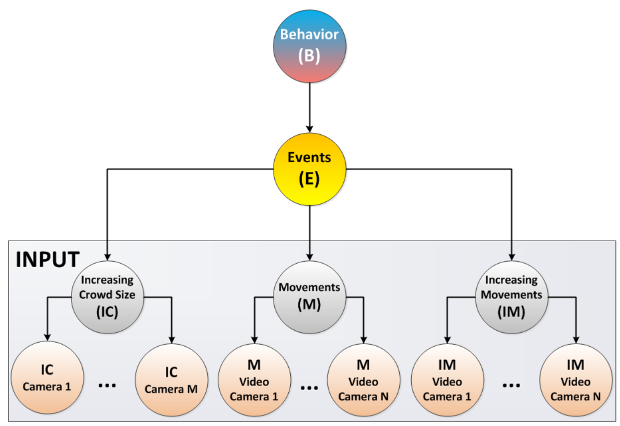


Figure 3: Bayesian model for the crowd behavior analysis. The highest level is the Behavior (B). This node depends on the previous node (Events), which are dependent from the Input, where the crowd analysis is computed.

Differently from the binary observation from the HMM input, the BN inputs (*In*) are divided into three discrete representations : Zero (*O*), Low (*L*) and High (*H*). The crowd analysis is decomposed in three kinds of data:

- Increasing Crowd Size (*IC*) - It is based in the crowd size explained in the Section 2.1. This input is generated using the variation in the crowd size considering the data in previous time. It is only used in airport scenarios in BN approach. The threshold used to Zero observation is smaller then three person. If the variation is less then five person the value of IC is *L*, else the value of IC is *H*.

- Movements (*M*) - This data is generated using optical flow information as explained in the Section 2.2. It is generated for three different cameras in the same scene in the ATM scenario and two cameras in airport scenarios. Threshold values used in this case are dependent of the camera and the scenario. Considering the observations from the HMM : The Zero occurs in the movements data if the observation O1 is true; if the observation O2 is true then the movements is Low (*L*). This data is High (*H*), if the observation O3 is true.

- Increasing Movements (*IM*) - This data is obtained using the variation in the movements data (*M*). Basically, threshold values are estimated as in the Movements data, and they are dependent of the camera and the scenario. This input data is related to observations O6 and O7 from HMM, but it is divided in three levels: *O*, *L* and *H*.

This Bayesian net fuses the data from different cameras using an uniform distribution, *i.e.* all data has the same importance and confidence. The distribution used can vary in other scenarios and applications. The fusion is achieved using three input nodes (*IC*, *M*, *IM*), shown in the Figure 3. The node Events (*E*) defines three possible events in the crowd behavior. Two are associated to normal behavior as **Calm Movements** (most of the people standing and few people walking), **Low Movements**: it is associated to crowd walking and interaction between persons. The **High Movements** is associated to strong movements in the crowd, it is related to abnormal behavior as fight, robbery, running, etc.

The learning step of the Bayesian net is an important limitation, since we do not have enough recorded training data. We derived training data based on knowledge and experience of the joint distribution of each variable in the Bayesian net, generating the Conditional Probability Tables (CPT).

The result of the Bayesian net uses the loglikelihood ratio test, it is used to compare the fit of two models one of which is nested within the other, in this case the normal and abnormal behavior. It is shown in the Equation 4.

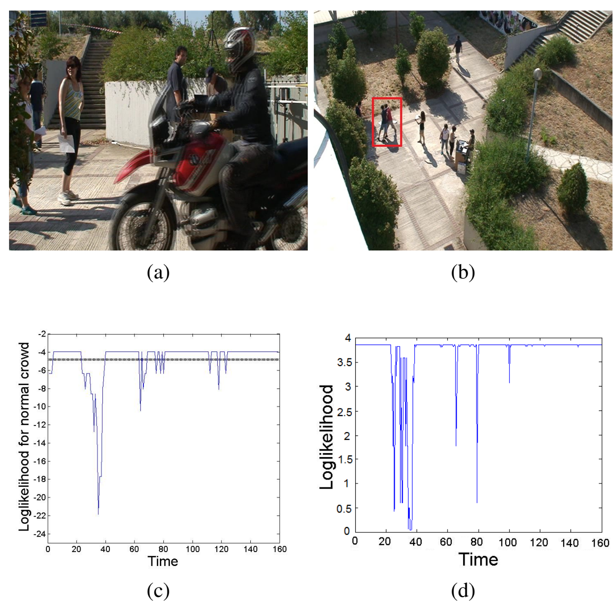
$$Loglikelihood = -\log \frac{P(B = Abnormal|E, In)}{P(B = Normal|E, In)}. \quad (4)$$

4 Experimental Results

We used four film sequences to illustrate abnormal behavior and compare the two proposal methods. The first three sequences have been acquired by two video cameras and one thermal camera in an outdoor scenario(airport) and fight behaviors occur in these film sequence. The normal behavior in this case is associated to a person walking to the queue, waiting to be attended, being attended and walking to the exit. In the last sequence, it has a robbery at an ATM machine. This sequence is composed by a set of four video cameras in different views, where three cameras are used to estimate the crowd activity and one to estimate crowd size. Frames for these filme sequences are shown in the Figure 7. In this scenario, there are a lot of people walking near the ATM, but only few persons stop in the ATM machine and interact with it.

The results are shown in the figures 4, 5, 6 and 7. All these figures show: the frames from the film sequence, the result generated by HMM approach and the result generated by BN approach. The results in y-axis shows the likelihood for normal behavior and in x-axis represents time in seconds. The grey straight line in HMM results indicates expected normal behavior.

The results of the crowd behavior analysis for Fight 1 sequence are shown in the Figure 4. A motorcycle is driving through the area at $20s <$ Time $< 40s$ (also an abnormal event in this case), it is illustrated in the Figure 4a where a frame from camera 1 is shown. The Fight 1 starts at Time $\sim 60s$ and ends at Time $\sim 100s$, it is illustrated in the Figure 4b where a frame from camera 2 is shown. In HMM results, there are some abnormal behaviors at Time $\sim 120s$ seconds, in this period two persons try to help the one lying down during the fight. It generates high movements but these movements increase very slow. Thus, the BN approach have not detected this abnormal behavior. In this scenario, the result of each approach is similar. But the “alarm” time duration in HMM approach is greater than in BN approach. In the BN approach results, the abnormal behavior occurs in a short time, but it is enough to ring an alarm.

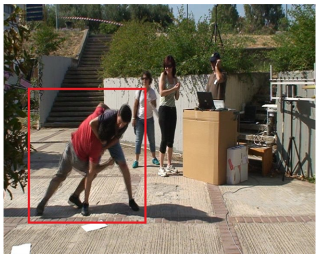


(a) (b) (c) (d)

Figure 4: Fight 1 sequence in outdoor scenario - a) Image from camera 1 at time $\sim 30s$ when a motorcycle is driving through the area; b) Image from camera 2 at Time $\sim 90s$ where a Fight occurs (red box); c) and d) the results by HMM and BN approach, respectively.

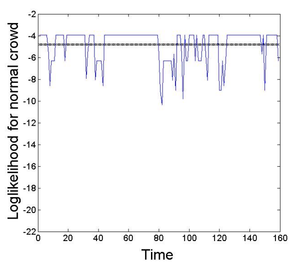
Fight 2 sequence is shown in the Figure 5. The fight occurs at Time $\sim 80s$ and ends at Time $\sim 120s$, it is illustrated in the Figure 5b,c, where two frame at Time $\sim 90s$ from camera 1 and 2 are shown, respectively. After that, the persons near fight help the one that lying down at $120s <$

Time $< 160s$. In HMM results, there are some abnormal behavior at Time $\sim 10s <$ Time $< \sim 50s$, they are false alarms. In the BN results, there are not false alarm if the abnormal threshold is set to value smaller than three. Nevertheless, the BN detects the abnormal events as fight and high movements during the help to the lying down person.

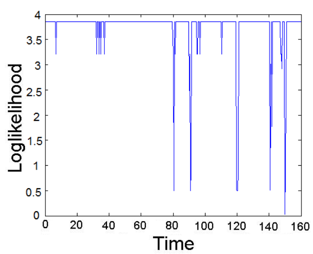


(a)

(b)



(c)



(d)

Figure 5: Fight 2 sequence in outdoor scenario - a) Image from camera 1 at Time $\sim 100s$ where a fight occurs, detailed using red boxes(Same moment in camera 2 is shown in b)); c) and d) results by HMM and BN approach, respectively.

Fight 3 starts at $\sim 120s$ and ends at $\sim 130s$ and it is illustrated in the Figure 6. In this case, fight occurs out of the camera 1 field of view. Due this difficulty, the results presented a set of false alarms. Both methods generated very similar results in this case, with a little difference in the beginning of the sequence, where the HMM approach detects a wrong abnormal behavior on the contrary of the BN approach.

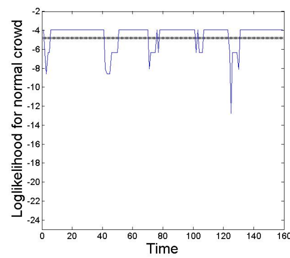
Robbery sequence is shown in the Figure 7. There are two persons close to the ATM in time $\sim 80s$. The robbery takes place at $\sim 120s$. After that, two persons (robber and robbery victim) running from ATM at 120s < Time < 160s . In HMM results, there are some abnormal behavior at Time $\sim 10s <$ Time $< \sim 60s$, they are false alarms. In the BN result, there are not false alarms, only in the Time $\sim 50sm$ when a shadow of one person appear very near camera 3. Nevertheless, the BN detects the abnormal events as robbery and running near the ATM machine, as well as the HMM.

5 Conclusions

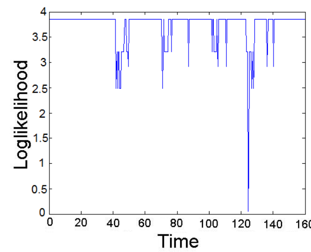
This paper describes a probabilistic approach to crowd behavior analysis. The information of crowd size and activity is computed in order to detect behaviors. The use of a network of sensor able the system to deal with abnormal behaviors. Two probabilistic methods are proposed to analyse crowd behavior. They are tested and compared in four different situation. Both the methods are able to detect abnormal behavior in crowd, with the advantage of the BN approach is less susceptible to false alarms. In the other hand, the HMM approach detects the abnormal behavior during a larger time period.

The method's capacity to detect abnormal behavior is strongly dependent on their parameters, and they are hard to estimate. The lack of datasets to training these methods limit the capable of both methods. Although, these methods are sensible to the parameters, they are able to detect abnormal behavior in a set of different scenarios.

(a)



(b)



(c)

Figure 6: Fight 3 sequence in outdoor scenario - a) Image from camera 2 at Time $\sim 125s$ where a Fight occurs (red box); c) and d) results by HMM and BN approach, respectively.

6 Future work

Future works will focus in test and evaluate the system with other experimental data representing a big set of abnormal behavior. The use of other information source as sound will be investigated. Moreover, the use of features in the scene that consider it as a whole is useful, but it could be improved if it uses an individual information of each person in the scene. The use of the Laban Movement Analysis [18] together with crowd analysis could improve the performance of the system as a whole. One possible approach is to use an hybrid probabilistic approach in order to get the best characteristics of each approach (HMM and BN).

The authors participate in the on-going EU funded project Prometheus (FP7-214901) [1]. Prometheus aims at establishing a general framework which links fundamental sensing tasks to automated cognition processes. The framework

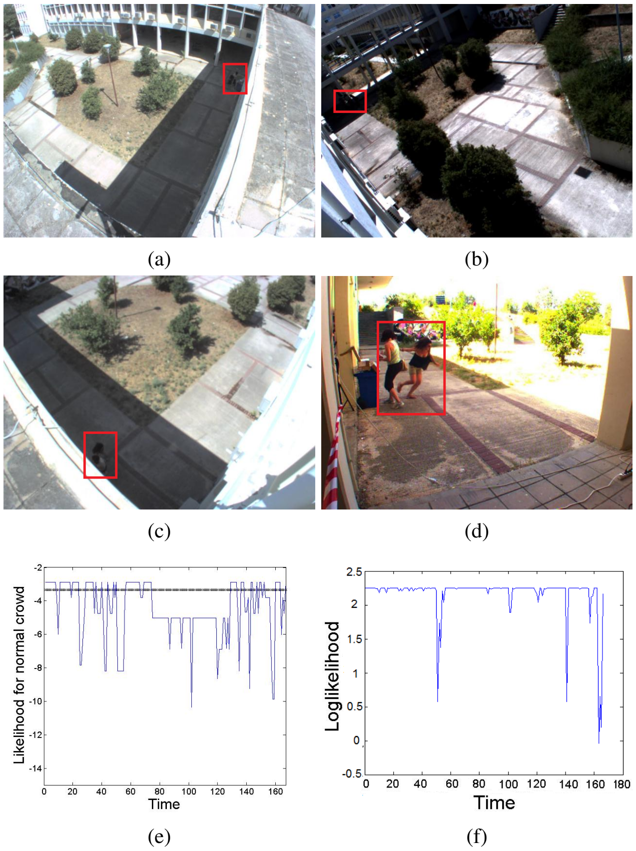


Figure 7: Robbery sequence in ATM scenario - a), b), c) and d) Image at the robbery in cameras 1,2,3 and 4, respectively (red box); e) and f) results by HMM and BN approach, respectively.

will enable interpretation and short-term prediction of individual and crowd behaviors. An important task is the definition and design of fusion models, tracking models and behavioral models that will be used to automatically detect persons and interpret their behavior as well as the behavior of groups of people. The work that has been presented in this article will be further developed in the continuing work of Prometheus.

7 Acknowledgments

This work has been supported by the European Union within the FP7 Project PROMETHEUS (EU-FP7-ICT-2007-1-214901), www.prometheus-FP7.eu.

References

[1] J. Ahlberg, D. Arsic, T. Ganchev, A. Linderhed, P. Menezes, S. Ntalampiras, T. Olma, I. Potamitis, and J. Ros. Prometheus: Prediction and interpretation of human behavior based on probabilistic structures and heterogeneous sensors. In *European Conference on Artificial Intelligence (ECAI)*, Patras, Greece, 2008.

[2] S. Ali and M. Shah. A lagrangian particle dynamics approach for crowd flow segmentation and stability analysis. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–6, 2007.

[3] S. Ali and M. Shah. Floor fields for tracking in high density crowd scenes. In *European Conference on Computer Vision (ECCV)*, pages 1–14, 2008.

[4] M. Andersson, J. Rydell, and J. Ahlberg. Estimation of crowd behavior using sensor networks and sensor fusion. In *International Conference on Information Fusion (FUSION '09)*, pages 396–403, 2009.

[5] E. Andrade, S. Blunsden, and R. Fisher. Modelling crowd scenes for event detection. In *International Conference on Pattern Recognition (ICPR)*, pages 175–178, Washington, DC, USA, 2006. IEEE Computer Society.

[6] H. Ardö. *Multi-target tracking using on-line Viterbi optimisation and stochastic Modelling*. PhD thesis, Mathematical Sciences - Lund, 2009.

[7] D. Arsic, J. Schenk, B. Schuller, F. Wallhoff, and G. Rigoll. Submotions for hidden markov model based dynamic facial action recognition. In *IEEE International Conference on Image Processing*, pages 673 – 676, 2006.

[8] G. Brostow and R. Cipolla. Unsupervised bayesian detection of independent motion in crowds. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 594–601, 2006.

[9] F. Cupillard, F. Bremond, and M. Thonnat. Behaviour recognition for individuals, groups of people and crowd. In *IEEE Symposium on Intelligence Distributed Surveillance Systems*, pages 7/1 – 7/5, 2003.

[10] D. Helbing and P. Molnar. Social force model for pedestrian dynamics. *Physical Review E*, 51:4282, 1995.

[11] D. Knill and A. Pouget. The bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27(12):712–719, 2004.

[12] N. MacFarlane and C. Schofield. Segmentation and tracking of piglets in images. *Machine Vision and Applications*, 8(3):187–193, 1995.

[13] J. Marques, P. Jorge, A. Abrantes, and J. Lemos. Tracking groups of pedestrians in video sequences. In *Conference on Computer Vision and Pattern Recognition Workshop (CVPRW)*, volume 9, pages 101–108, 2003.

[14] R. Mehran, A. Oyama, and M. Shah. Abnormal crowd behavior detection using social force model. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 935–942, Los Alamitos, CA, USA, 2009. IEEE Computer Society.

[15] N. Pelechano and N. Allbeck, J.and Badler. Controlling individual agents in high-density crowd simulation. In *ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA 07)*, pages 99–108, Aire-la-Ville, Switzerland, Switzerland, 2007. Eurographics Association.

[16] L. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Readings in speech recognition*, pages 267–296, 1990.

[17] P. Reisman, O. Mano, S. Avidan, and A. Shashua. Crowd detection in video sequences. In *IEEE Intelligent Vehicles Symposium (IV2004)*, pages 66–71, 2004.

[18] J. Rett, J. Dias, and J. Ahuactzin. Bayesian reasoning for laban movement analysis used in human machine interaction. *Int. J. Reasoning based System*, 1:64–74, 2008.

[19] B. Zhan, D. Monekosso, P. Remagnino, S. Velastin, and L. Xu. Crowd analysis: a survey. *Machine Vision Application*, 19(5-6):345–357, 2008.

[20] T. Zhao and R. Nevatia. Tracking multiple humans in complex situations. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 26(9):1208–1221, 2004.