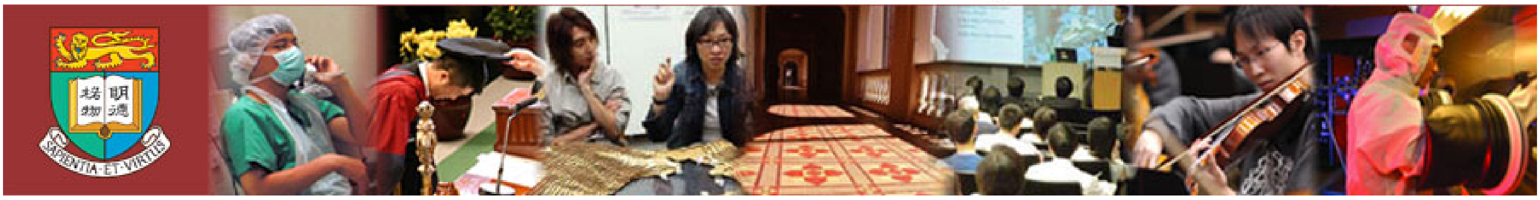

A Self-Growing Bayesian Network Classifier for Online Learning of Human Motion Patterns

Zhuo Chen   Nelson H.C. Yung
Department of Electrical and Electronic Engineering
The University of Hong Kong
Pokfulam Road, Hong Kong, H.K.S.A.R.
zchen@eee.hku.hk   nyung@eee.hku.hk

Abstract—**This paper proposes a new self-growing Bayesian network classifier for online learning of human motion patterns (HMPs) in dynamically changing environments. The proposed classifier is designed to represent HMP classes based on a set of historical trajectories labeled by unsupervised clustering. It then assigns HMP class labels to current trajectories. Parameters of the proposed classifier are recalculated based on the augmented dataset of labeled trajectories and all HMP classes are accordingly updated. As such, the proposed classifier allows current trajectories to form new HMP classes when they are sufficiently different from existing HMP classes. The performance of the proposed classifier is evaluated by a set of real-world data. The results show that the proposed classifier effectively learns new HMP classes from current trajectories in an online manner.**

Keywords-Bayesian network classifier; online learning; human motion patterns

I. INTRODUCTION

A classifier plays a significant role in data analysis and pattern recognition. It can be defined as a function that assigns a class label to instances based on a set of attributes by a classification task, of which patterns of data are learned from the dataset of labeled instances. Thus, the induction of a classifier is an important problem in the field of pattern recognition, machine learning and data mining.

Naive Bayesian classifier can be considered as one of the most effective classifiers in the sense that it often works quite well in practice [1]. It has been widely applied for resolving different practical problems [2-4]. This classifier learns the conditional probability of each attribute given the class label. Classification is then performed by selecting the class with the highest posterior probability of the class given the particular instance that is computed by the Bayes rule. This computation is made based on a strong assumption that all the attributes of the instance are conditionally independent. Although naive Bayesian classifier has a competitive performance, other classifiers have been developed for better performance by relaxing the restrictive assumption since it may not be realistic in practical cases. This motivates the development of Bayesian network classifiers.

Bayesian network is a probabilistic model that describes a set of attributes and their conditional dependences by using a directed acyclic graph. In such a graph, each node represents an attribute and directed edges represent conditional dependences between attributes. When an instance described by a set of attributes is classified, a Bayesian network classifier considers the conditional dependences between attributes. As a result, Bayesian network classifier has been used in a wide range of applications [5-8]. Despite an improved idea and a better classification performance, Bayesian network classifier has a number of limitations. When classifying an instance, the classifier provides the best choice among all existing classes based on the highest posterior probability, but in some case, the instance may not be similar to any of the existing classes. In other words, the instance likely represents a new class rather than one of the existing classes. For example, when classification of trajectories is performed for online learning human motion patterns (HMPs), a current trajectory may be quite different from any existing HMP class thus should be regarded as a new HMP class. If a basic Bayesian network classifier is used, the current trajectory would be incorrectly classified to an existing HMP class with the highest probability. Thus, we intend to provide a solution to this problem.

In this paper, we propose a new self-growing Bayesian network classifier for online learning HMPs by classifying current trajectories. A set of recorded historical trajectories are first labeled by using an unsupervised clustering method. The process of designing the proposed classifier can be then described as building a probability model for learning each existing HMP based on the labeled historical trajectories in each class. When current trajectories are constantly obtained from the dynamically changing environment, each of them is classified into an existing HMP class or being considered as a new HMP class. By passing each classified trajectory to the previous set of labeled trajectories, parameters of the probability model are accordingly recalculated and the proposed classifier is self-grown which can represent the updated existing HMPs and new born HMPs in an online manner. Thus, in the process of performing classification, the proposed classifier not only has traditional function of a classifier for classifying the current trajectory to the most similar HMP class but also is capable of detecting new HMP class when the current trajectory is quite different from all existing HMP classes. A set of real-world experiments are presented to test the performance of the self-growing Bayesian network classifier for online learning HMP classes. The results show that the proposed classifier is effective for both updating existing HMP classes and detecting new HMP classes.

The rest of this paper is organized as follows. In Section II, the generalized framework for online learning of HMP classes is presented. Section III introduces the proposed self-growing Bayesian network classifier. Section IV depicts the results of a set of real-world experiments performed by the proposed classifier. In Section V, the whole paper is concluded and future research work is discussed.

II. Generalized Framework for Online Learning of Human Motion Patterns

The generalized framework for online learning of HMP classes consists of four main functions: (1) Trajectory Formation; (2) Trajectory Clustering; (3) HMP Learning; (4) Trajectory Classification; as depicted in Figure 1. Human trajectories are extracted from the video in Trajectory Formation module, in which many methods for human detection and tracking can be used. Here, we choose a method of using foreground extraction and model-fitting in a single frame and data association across frames [9, 10] for our purpose.

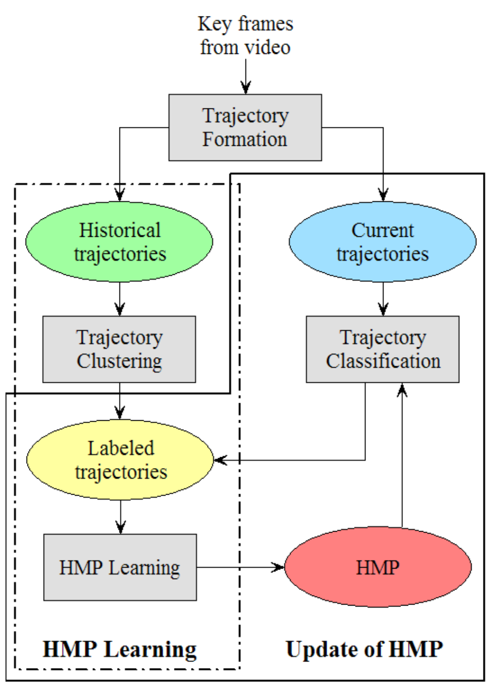

Figure 1. Generalized framework for online learning of HMP classes.

Based on extracted trajectories, the whole online learning process of HMP classes can be separated into two stages: HMP Learning and Update of HMP. In the stage of HMP Learning, the derived historical trajectories are first clustered in Trajectory Clustering module, by using a constrained gravitational clustering approach which was presented in our previous research work [11]. Based on the obtained labeled trajectories, a self-growing Bayesian network classifier is built and HMP classes are learned and described by a probability model in Motion Pattern Learning module.

The stage of Update of HMP is an iterative process which is automatically performed once a current trajectory is derived. In Trajectory Classification module, the current trajectory is first classified into a HMP class by the self-growing Bayesian network classifier. With the current trajectory changing to be a labeled trajectory, parameters of the classifier are recalculated based on the updated labeled trajectories dataset. The mean vector and the prior probability of each HMP class are also updated. Thus, all the HMP classes can be online learned according to the dynamic change in the environment.

III. Self-Growing Bayesian Network Classifier

Based on the generalized framework described in Section II, the focus of the online learning of HMP classes is the self-growing Bayesian network classifier. In this section, we discuss the details of the proposed classifier through the following specifications of attributes of interest for classification, structure of the classifier, and learning of the classifier.

A. Attributes of Interest for Classification

For representing a human trajectory, we consider spatial location, velocity and change of heading angle as attributes of interest when classifying current trajectories. We define a feature vector to incorporate these attributes of interest, which aims to describe the motion characteristic of human. Let $\mathbf{T}_k$ denotes the feature vector of the k^{th} human trajectory. Based on the attributes of interest we consider here, $\mathbf{T}_k$ is given as:

$$\mathbf{T}_k = \{o_x^{(k)}, o_y^{(k)}, d_x^{(k)}, d_y^{(k)}, v_m^{(k)}, v_d^{(k)}, c^{(k)}, n_l^{(k)}, n_r^{(k)}\}, \quad (1)$$

where $(o_x^{(k)}, o_y^{(k)})$ denotes the origin of the trajectory, and $(d_x^{(k)}, d_y^{(k)})$ denotes the destination of the trajectory. Since there are different velocity value at each time step, $v_m^{(k)}$ and $v_d^{(k)}$ denote the mean and the standard deviation of velocity value at all time steps of the trajectory, respectively. $c^{(k)}$ denotes the curvature of trajectory. $n_l^{(k)}$ and $n_r^{(k)}$ denote the number of time steps of turning left and right when compared with the moving direction at the previous time step, respectively.

B. Structure of the Classifier

The self-growing Bayesian network classifier is designed based on a matching mechanism between the current trajectory and each HMP class and a decision function operated on the corresponding matching results, as depicted in Figure 2.

When matching the current trajectory with each HMP class, the difference and the similarity between the current trajectory and the HMP class are respectively measured in hierarchical levels. A negative operator first decides whether the trajectory is sufficiently different from the

HMP class to be a negative case of the HMP class, meaning that the trajectory does not belong to the HMP class. If the trajectory is not regarded as a negative case of the HMP class by the negative operator, a positive operator further calculates a likelihood value for that the trajectory belongs to the HMP class. The matching mechanism generates the results of measuring difference and similarity between the trajectory and each HMP class, respectively. A decision function is finally used to assign a class label of HMP for the current trajectory.

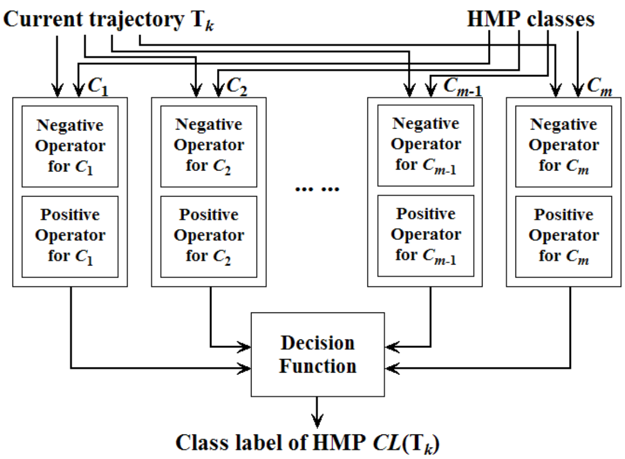

Figure 2. Structure of self-growing Bayesian network classifier.

C. Learning of the Classifier

For the negative operator in the classifier, we proposed a function $NC(\mathbf{T}_k, C_i)$ for judging whether the current trajectory $\mathbf{T}_k$ is a negative case of the HMP class C_i. $NC(\mathbf{T}_k, C_i)$ can take the form of a step function or a Gaussian function. For example, the step function form of $NC(\mathbf{T}_k, C_i)$ is given as:

$$NC(\mathbf{T}_k, C_i) = \begin{cases} 1, & D(\mathbf{T}_k, C_i) > D^{(B)} \\ 0, & D(\mathbf{T}_k, C_i) \le D^{(B)} \end{cases}, \quad (2)$$

where $D(\mathbf{T}_k, C_i)$ is the Euclidean distance between $\mathbf{T}_k$ and the centroid of C_i which is represented by the mean vector of all trajectories in C_i, and $D^{(B)}$ is a boundary distance which refers that two feature vectors are likely to fall into the same class only when $D(\mathbf{T}_k, C_i)$ is less than or equals to $D^{(B)}$. Thus, when $D(\mathbf{T}_k, C_i)$ is larger than $D^{(B)}$, the function $NC(\mathbf{T}_k, C_i)$ returns “1” that means the trajectory $\mathbf{T}_k$ is a negative case of the HMP class C_i.

The criterion for determining $D^{(B)}$ is based on the distance distribution of all pairs of labeled trajectories which is give as:

$$D^{(B)} = \frac{D_{\max\_freq} + D_{\min\_freq}}{2}, \quad (3)$$

where D_{max_freq} and D_{min_freq} are the distance values respectively corresponding to the first local maximal and minimal frequency values in the distance distribution as depicted in Figure 3. We evenly divide the range of distance values into 20 bins. The distance distribution is then obtained by calculating the average of all distance value that fall into a bin as the representative distance value of the bin, and counting the number of all these distance value falling into the bin as the corresponding frequency value for the bin. When performing unsupervised clustering for labeling historical trajectories in Trajectory Clustering module, we observe that (1) [D_{max_freq}, D_{min_freq}] is an important range obtaining optimal clustering; and (2) Equation (3) can decide an effective adaptive threshold for representing the range.

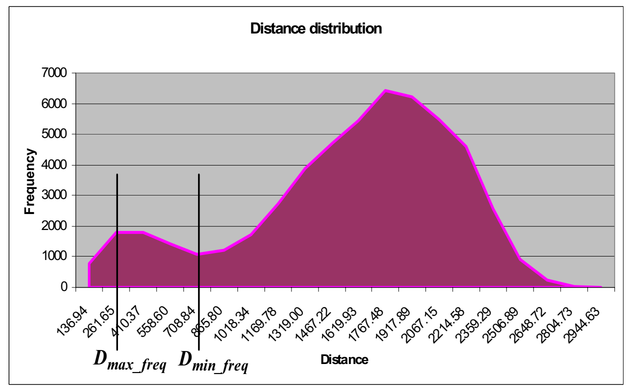

Figure 3. D_{max_freq} and D_{min_freq} in the distance distribution.

For the positive operator in the classifier, we use a *tree-augmented naive Bayesian* (TAN) network proposed in [12] to capture conditional dependences between attributes in the feature vector of human trajectory mentioned in Section III.A. Since it is NP-complete to find all the conditional dependences between attributes in the feature vector to build a complete Bayesian network, some restriction is necessary to make the idea of Bayesian network applicable in solving practical problems. TAN network is a kind of Bayesian network that assumes each attribute has conditional dependence on at most one other attribute in the feature vector. The use of TAN network results in a balance between improving the performance of naive Bayesian classifier and simplifying the NP-complete problem of searching all the conditional dependences to be efficiently resolvable. The algorithmic steps of learning TAN network for representing correlations between attributes in the feature vector of human trajectory are given as:

Step 1: Compute mutual information $I(A_j, A_k)$ between each pair of attributes (A_j, A_k) $(j \neq k)$ in the feature vector, where

$$I(A_j, A_k) = \sum_{i=1}^{m} P(C_i) \log\left(\frac{1}{\sqrt{1-\rho_{(jk)|C_i}^2}}\right). \quad (4)$$

In our problem, we assume each attribute in the feature vector is a Gaussian, so $\rho_{(jk)|C_i}$ means the correlation coefficient between A_j and A_k given the HMP class C_i. $P(C_i)$ is prior probability of C_i, and m is the number of all existing HMP classes.

Step 2: Build a complete undirected graph in which each node is an attribute in the feature vector and the weight of each edge is the value of mutual information computed in *Step 1*.

Step 3: Find a maximum weighted spanning tree of the built complete graph.

Step 4: Transform the obtained undirected spanning tree to a directed one by randomly choosing a root node and setting the direction of all edges away from the root.

Based on the learned TAN network, the joint probability distribution function over the whole feature vector $\mathbf{T}_k$ of the human trajectory given the HMP class C_i can be written as:

$$p(\mathbf{T}_k \mid C_i) = p(A_1, A_2, ..., A_9 \mid C_i) = \prod_{j=1}^{9} p(A_j \mid \Gamma_{A_j}, C_i), \tag{5}$$

where Γ_{A_j} represents the parent node of A_j in the learned TAN network which means A_j has a direct dependence on Γ_{A_j}. Here, "9" means the number of attributes which are defined in Equation (1). A likelihood $L(C_i|\mathbf{T}_k)$ is then calculated for measuring the possibility that $\mathbf{T}_k$ belongs to C_i. The calculation of $L(C_i|\mathbf{T}_k)$ is given as:

$$L(C_i \mid \mathbf{T}_k) = P(C_i) \cdot p(\mathbf{T}_k \mid C_i), \tag{6}$$

where $P(C_i)$ is the prior probability of C_i which is calculated as the ratio of the number of trajectories in the HMP class C_i and the number of all labeled trajectories.

After performing the negative and positive operators for each HMP class C_i, we proposed a decision function $CL(\mathbf{T}_k)$ for generating the class label of HMP for the current trajectory $\mathbf{T}_k$ which is given as:

$$CL(\mathbf{T}_k) = \begin{cases} i: & \exists C_i (1 \le i \le m), NC(\mathbf{T}_k, C_i) = 0 \wedge \arg\max_i \{L(C_i \mid \mathbf{T}_k)\} \\ m+1: & \forall C_i (1 \le i \le m), NC(\mathbf{T}_k, C_i) = 1 \end{cases}. \tag{7}$$

IV. EXPERIMENT

Figure 4. The scenario of the real-world experiment.

In this section, we demonstrate how the proposed classifier works for online learning HMP classes. The scenario of the experiment is based on people walking freely in a shopping mall as shown in Figure 4, which is a representative scenario of many generic real-world environments. There are several entrances and exits to the scene as depicted: entrance/exit '1' connects to more shops; entrances/exit '2' connects to neighboring buildings; entrance/exit '3' connects to up/down escalators. A background-fixed video was taken for the scenario and a total of 326 human trajectories were accordingly extracted as shown in Figure 5, in which red curves and green curves represent bi-directional trajectories between each pair of entrance and exit.


Figure 5. Extracted human trajectories.

TABLE I. DESCRIPTION OF CURRENT TRAJECTORIES USED IN THE EXPERIMENTS

	Dataset of Current Trajectories	
	# Trajectories from existing HMPs	***# Trajectory from new HMPs***
1	51	9
2	4	56
3	2	58
4	41	19
5	20	40
6	45	15
7	28	32
8	11	49
9	58	2
10	40	20
11	46	14
12	59	1
13	58	2
14	56	4
15	58	2
16	59	1
17	59	1
18	59	1
19	48	12
20	45	15
Average description	42.35	17.65

In order to simulate and test the online learning process of HMPs based on the proposed classifier, we separate all 326 human trajectories into two groups. We randomly select 266 extracted trajectories as historical trajectories for building the proposed classifier, and regard the remaining 60 trajectories as current trajectories for performing classification and updating HMPs. For evaluation of the robustness of the proposed classifier, we perform 20 experiments by randomly separating all extracted human trajectories to group the datasets of historical trajectories and current trajectories, respectively. In order to test the special property of the proposed classifier that to be superior to traditional classifiers in detecting new classes, we avoid that testing current trajectories all come from the existing HMPs. The description of current trajectories for classification which are used in total 20 experiments is depicted in Table I. It can be seen that there are current trajectories from new HMPs in each experiment and in some experiments even more current trajectories are from new HMPs than existing HMPs. On average, nearly 30% current trajectories are from new HMPs which is a reasonable proportion for testing the performance of detecting new HMPs of the proposed classifier.

TABLE II. EXPERIMENTAL RESULTS BY THE PROPOSED CLASSIFIER

	# Trajectories of incorrect classification	*Accuracy rate*
1	1	98.33%
2	7	88.33%
3	12	80%
4	4	93.33%
5	1	98.33%
6	1	98.33%
7	3	95%
8	8	86.67%
9	1	98.33%
10	2	96.67%
11	3	95%
12	3	95%
13	1	98.33%
14	7	88.33%
15	0	100%
16	2	96.67%
17	0	100%
18	4	93.33%
19	5	91.67%
20	6	90%
Average performance	3.55	94.08%

Since all 326 human trajectories are obtained from a dynamically changing environment in which no prior knowledge of HMPs is available, no objectively correct information of HMP class label for each trajectory can be derived. In order to analyze the accuracy of the proposed classifier, we first perform Trajectory Clustering for labeling human trajectories, and the corresponding class label information of the human trajectories are used as a referenced truth for comparison with the classification results by the proposed classifier. The analysis of classification results of current trajectories by the proposed classifier is given in Table II. Among 60 current trajectories in each experiment for testing the classifier, we could see that there is no incorrect classification when the classifier has the best performance and 12 current trajectories are incorrectly classified when the classifier has the worst performance. Globally observing the results of all 20 experiments, the proposed classifier could obtain 94.08% accuracy rate in average when classifying the current trajectory into some existing HMP class or developing it as a new HMP class.

Figure 6. Correct detection of new HMP by the proposed classifier.

Figure 7. Failed case of classification by the proposed classifier.

For illustration purpose, a successful case and a failed case of classification are shown in Figure 6 and Figure 7, respectively. As shown in Figure 6, the blue curve depicts a current trajectory which is successfully considered as a new HMP by the proposed classifier rather than incorrectly classified to the most similar existing HMP which is represented by the red curve. The circles label the destination information of the classified trajectory and the HMP. The green curves show the corresponding historical trajectories in the HMP, and it can be clearly seen that they are quite different from the current trajectory. However, the proposed classifier may fail in some boundary case, as depicted in Figure 7. The blue curve represents a current trajectory which should be reasonably classified to an existing HMP shown as the red curve since the current trajectory is quite similar to some labeled historical trajectories in the HMP represented by the green curves, while the proposed classifier regards the current trajectory as a negative case of the similar existing HMP because of a relatively large distance between them. Thus, a better measurement of difference between the current trajectory and the HMP in the negative operator need be explored that would improve the performance of the proposed classifier. The circles also label the corresponding destination information.

V. CONCLUSION

In this paper, we presented a new self-growing Bayesian network classifier for online learning HMPs in dynamically changing environments. The proposed classifier is built for describing HMPs based on a set of labeled trajectories obtained from unsupervised clustering process, and used for assigning the class label of HMP to current trajectories. The proposed classifier is then rebuilt based on the augmented dataset of labeled trajectories and all HMP classes are accordingly updated. The main advantage of this new self-growing Bayesian network classifier concentrates on that the classifier can not only classify the current trajectory into an existing HMP class with the highest similarity score but also assign a new class label to the current trajectory when it is quite different with all existing HMP classes. A set of real-world experiments are performed for online learning HMPs in a dynamically changing environment in which the performance of the self-growing Bayesian network classifier is tested. The results show that proper classification of current trajectories are obtained by using the proposed classifier, based on which both existing HMPs and new HMPs can be effectively online learned according to the dynamic change in the environment. Based on the generalized online learning framework of HMPs by using the proposed classifier, our future research will focus on two issues: (1) to improve the online learning of HMPs more efficiently and accurately; (2) to investigate human behavior prediction based on the online learned HMPs.

REFERENCES

[1] R. O. Duda, P. E. Hart and D. G. Stork, Pattern classification, 2nd ed., Wiley-Interscience, New York, 2001.

[2] I. Colak, M. Demirtas, G. Bal and H.T. Kahraman, “A Parameter Determination System for Wind Turbines Based on Naive Bayes Classification Algorithm,” Proceedings of the Eighth International Conference on Machine Learning and Applications, Miami Beach, Florida, USA, December 2009.

[3] T.T. Huang and C.H. Chang, “Intelligent Stock Selecting via Bayesian Naive Classifiers on the Hybrid Use of Scientific and Humane Attributes,” Proceedings of the Eighth International Conference on Intelligent Systems Design and Applications, Kaohsiung City, Taiwan, November 2008.

[4] K.B. Wagholikar, S. Vijayraghavan and A.W. Deshpande, “Fuzzy Naive Bayesian Model for Medical Diagnostic Decision Support,” Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Minneapolis, Minnesota, USA, September 2009.

[5] S.S. Huang, L.C. Fu and P.Y. Hsiao, “Region-Level Motion-Based Foreground Segmentation Under a Bayesian Network,” IEEE Transactions on Circuits and Systems for Video Technology, Vol. 19, No. 4, pp. 522-532, 2009.

[6] M. Kim and V. Pavlovic, “Discriminative Learning of Mixture of Bayesian Network Classifier for Sequence Classification,” Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York, NY, USA, June 2006.

[7] K. Ueno, T. Hayashi, K. Iwata, N. Honda, Y. Kitahara and T.K. Paul. “Prioritizing Health Promotion Plans with *k-Bayesian Network Classifier*,” Proceedings of the Seventh International Conference on Machine Learning and Applications, San Diego, California, USA, December 2008.

[8] F. Lefevre, “Dynamic Bayesian Networks and Discriminative Classifiers for Multi-Stage Semantic Interpretation,” Proceedings of the 2007 International Conference on Acoustics, Speech, and Signal Processing, Honolulu, Hawaii, USA, April 2007.

[9] L. Wang and N. H. C. Yung, “Crowd Counting and Segmentation in Visual Surveillance,” Proceeding of the IEEE International Conference on Image Processing, Cairo, Egypt, November 2009.

[10] L. Wang and N. H. C. Yung, "Extraction of moving objects from their background based on multiple adaptive thresholds and boundary evaluation," IEEE Transactions on Intelligent Transportation Systems, vol. 11, no. 1, pp 40-51, 2010.

[11] Z. Chen, D.C.K. Ngai and N.H.C. Yung, “Pedestrian behavior prediction based on motion patterns for vehicle-to-pedestrian collision avoidance,” Proceedings of the 11th IEEE International Conference on Intelligent Transportation Systems, Beijing, China, October 2008.

[12] N. Friedman, D. Geiger and M. Goldszmidt, “Bayesian Network Classifiers,” Machine Learning, Vol. 29, No. 2, pp. 131-163, 1997.