

Pedestrian Behavior Prediction based on Motion Patterns for Vehicle-to-Pedestrian Collision Avoidance

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

D. C. K. Ngai
R&D Department
ASM Assembly Automation Ltd.
H.K.S.A.R.
aeeckngai@aaa.asmpt.com

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

Abstract— This paper proposes a prediction method for vehicle-to-pedestrian collision avoidance, which learns and then predicts pedestrian behaviors as their motion instances are being observed. During learning, known trajectories are clustered to form Motion Patterns (MP), which become knowledge a priori to a multi-level prediction model that predicts long-term or short-term pedestrian behaviors. Simulation results show that it works well in a complex structured environment and the prediction is consistent with actual behaviors.

I. INTRODUCTION

Vehicle-to-pedestrian accidents are a major concern for road safety in general. Compared with vehicles, pedestrians are physically more vulnerable, and suffer higher risk of death per kilometer traveled [1]. For many years, the number of pedestrian accidents remains at a plateau and shows no signs of reduction. It is clearly desirable to bring this number down [2]. Obviously, if potential collisions can be predicted, some of the vehicle-to-pedestrian accidents may be avoided altogether. Thus, collision avoidance has been an important research area for reducing pedestrian fatalities.

In a Dynamically Changing Environment (DCE) involving a vehicle and multiple pedestrians, the basic Collision Avoidance (CA) is to predict pedestrian behavior and decide a proper path for the vehicle itself. Conventionally, pedestrian behavior refers to pedestrian motion in the next time-step [3-8]. Techniques such as neural network [3], Markov models [4], Kalman filter [5, 6], and collision/velocity cones [7, 8] use current and historical motion data as their basis to predict motion in the next time step, which is short term and restrictive because they treat the CA problem locally and sub-optimally [9].

In order to perform CA more effectively, some researchers have recently attempted long-term behavior prediction for global CA [10, 11]. They predict over a number of future time-steps giving the vehicle a better chance of making the right action decision. Clearly, long-term behavior prediction is superior for avoiding collisions with globally optimal paths, but is naturally far more challenging. In [10], it treats the final destination

point of the pedestrian's movement as a long-term prediction goal, though this does not guarantee collision-free motion because there are many possible routes and motion patterns between each origin-destination pair. In [11], long-term prediction is made based on a set of trajectories between a number of resting places where people stop and stay. It requires the locations of these resting places be known a priori for the formation of Motion Patterns (MP) of the pedestrian. In more realistic DCE, prior knowledge of MP is usually not available. This generic scenario is what motivates our research.

In this paper, we propose a new prediction method that learns and predicts pedestrian behaviors as their motion instances are being observed. The observed trajectories (in terms of spatial location, velocity or heading angle) are clustered using the CGC algorithm [13] to form MP. For each clustered MP, it is evaluated for completeness against a criterion. They are then classified as *complete* MP (MP-C), which represents pattern that does not change much over time, or as *incomplete* MP (MP-I), which may be updated after subsequent prediction. Based on these MP, a multi-level prediction model is proposed. It consists of three levels of prediction in which the high and middle levels are both long-term predictions based on the MP-C and MP-I that predict future trajectories over a number of time-steps. The low level uses an AR model [14] to predict future trajectories over the next time step. The proposed method has been implemented and a number of simulation experiments have been conducted. The results show that it works well in a multiple pedestrian environment and the prediction is consistent with the actual behavior.

The rest of this paper is organized as follows. In Section II, we present an overview of the proposed method and a description of its main functions. Section III depicts the simulation results produced by the proposed method. Section IV concludes the paper with a brief discussion of future research direction.

II. PROPOSED METHOD

A. Overview

Our proposed method consists of four main functions: (1) Trajectory Formation; (2) MP Clustering; (3) MP

Classification; and (4) Pedestrian Behavior Prediction; as depicted in Fig. 1. In this paper, we assume that (a) the proposed method runs in a vehicle, which is to perform CA and other tasks in a DCE; (b) the method for determining the vehicle's action decision is outside the scope of this paper (Interested reader should refer to [12] for details); (c) an observable pedestrian instance is a distance value measured at time t by the vehicle's sensors; and (d) pedestrian behavior is defined by a series of pedestrian motions.

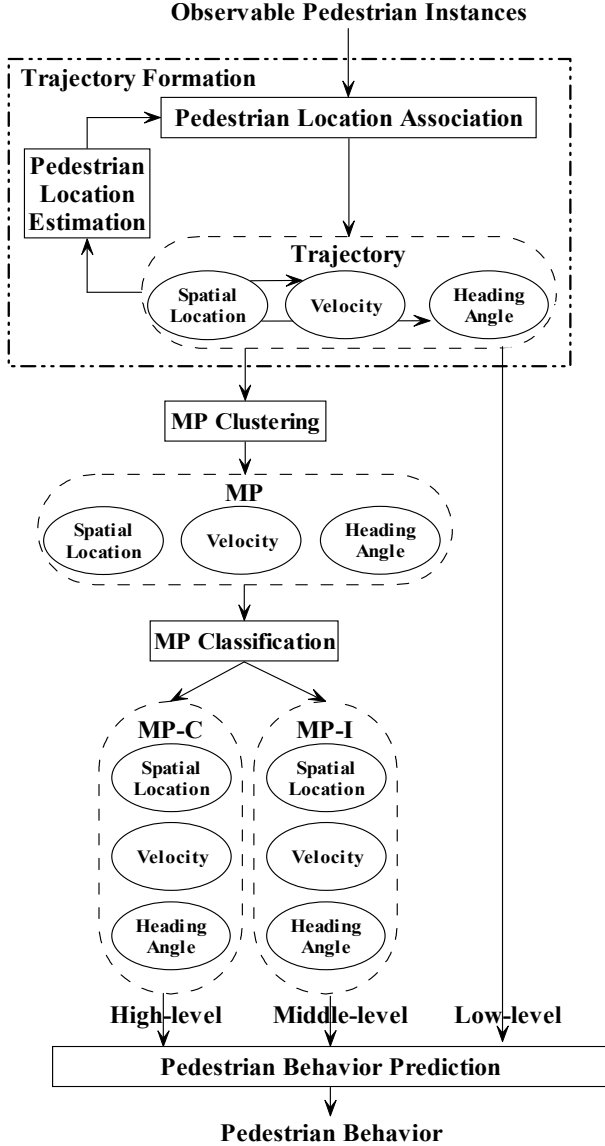


Fig. 1: Block diagram of the proposed method

To start with, when a group of pedestrian instances at time t are observed, they are first associated with existing trajectories that have been assembled through the previous $t-1$ time steps. The association relies on the $t-1$ instances to optimize a global shortest distance for all instances. From the newly formed trajectories with spatial location feature, we can derive the corresponding velocity and heading angle features. Given the trajectories, MP are clustered using an

instance-based clustering algorithm as a general representation of a sub-group of trajectories. Trajectories that do not cluster with others are only kept for one more time step, just in case they represent the birth of a new trajectory. Each clustered MP is further classified into MP-C or MP-I which is conditioned upon that the MP has been obtained based on a relative number of observable motion instances. Those MP that are evaluated as MP-C will be used for all subsequent predictions, and the MP-I may be modified or discarded after the behavior prediction is completed.

Finally, pedestrian behavior prediction accepts three inputs: (1) MP-C; (2) MP-I; and (3) current trajectories. Given a trajectory at time t , it is matched with the available MP-C first which is called high-level prediction if there is a close match between a trajectory and an MP-C, then the long-term behavior of the trajectory is predicted to be similar to the MP-C. If no MP-C is available or the match is negative, the remaining unmatched trajectories are matched with the MP-I. This is called middle-level prediction. If there is a match, then the medium-term behavior of a trajectory is predicted to be similar to the span of the matched MP-I. If all else fail, then prediction is performed for $t+1$ step using an Auto-Regressive (AR) model [14].

B. Trajectory Formation

In trajectory formation, we assume that the location information $(x_k(t), y_k(t))$ of pedestrian PD_k can be derived from distance sensors in the form of discrete time location information $\mathbf{r}_k[n]$ where $\mathbf{r}_k[n] = (x_k[n], y_k[n])$ given that $x_k[n] = x_k(nT_s)$ and $y_k[n] = y_k(nT_s)$ where T_s is the sampling time. We further assume that pedestrians cannot be identified otherwise such that we only know a set of sequential location information $\{\mathbf{r}_k[n]\}$ without the knowledge of which location information is associated with which pedestrian. The trajectory \mathbf{T}_k^s of PD_k for spatial location is defined by the set of sequential states associated with the pedestrian such that $\mathbf{T}_k^s = \mathbf{t}_k^s(n_1, n_2) = \{\mathbf{r}_k[n]\}$, where $n_1 \leq n \leq n_2$. We propose a candidate solution using an AR model to predict the possible current location of a pedestrian given its historical locations. We then associate the newly observed location to \mathbf{T}_k^s such that it is closest to the predicted location of \mathbf{T}_k^s .

Apart from spatial locations, other motion features such as velocity $\mathbf{v}_k[n]$ and heading angle $\phi_k[n]$ can also be derived and used for prediction at a later stage to give predicted velocity and heading angle profiles. The advantages of including these extra feature dimensions are that firstly, each pattern can be more comprehensively described, and secondly, consistency of patterns can be identified based on multiple dimensions depending on the need. In summary, PD_i is described by $\{\mathbf{T}_k^s, \mathbf{T}_k^v, \mathbf{T}_k^\phi\}$, where $\mathbf{T}_k^v = \mathbf{t}_k^v(n_1, n_2) = \{\mathbf{v}_k[n]\}$ and $\mathbf{T}_k^\phi = \mathbf{t}_k^\phi(n_1, n_2) = \{\phi_k[n]\}$ where $n_1 \leq n \leq n_2$.

C. MP Clustering

To cluster MP from trajectories of one of the three feature

dimensions, we employ the constrained gravitational clustering (CGC) method as described in [13]. The CGC method imposes a clustering constraint per iteration to control the formation of multiple clusters, without needing to assign a termination condition. At the start, each trajectory is regarded as the initial mean location vector of a cluster. In principle, the clustering process is completely controlled by the attraction and repulsion between existing clusters in the feature space. Analogy to gravitational force, existing clusters separated by a short distance are more likely to form a new cluster compared with those separated by a long distance. The ‘gravitational force’, $F_{T_k T_j}$ between the k^{th} and j^{th} clusters is given as:

$$F_{T_k T_j} = G \frac{m_k \times m_j}{|l_k - l_j|^3} (l_k - l_j), \quad (1)$$

where G is the gravitational constant which is set to 1 here, m_k and m_j are the masses represented by the numbers of trajectories in the k^{th} and j^{th} clusters respectively, and l_k and l_j are the mean location vectors in the feature space. Similar clustering effect can be found in the velocity and heading angle trajectories. The MP clustering step is repeated at each time step.

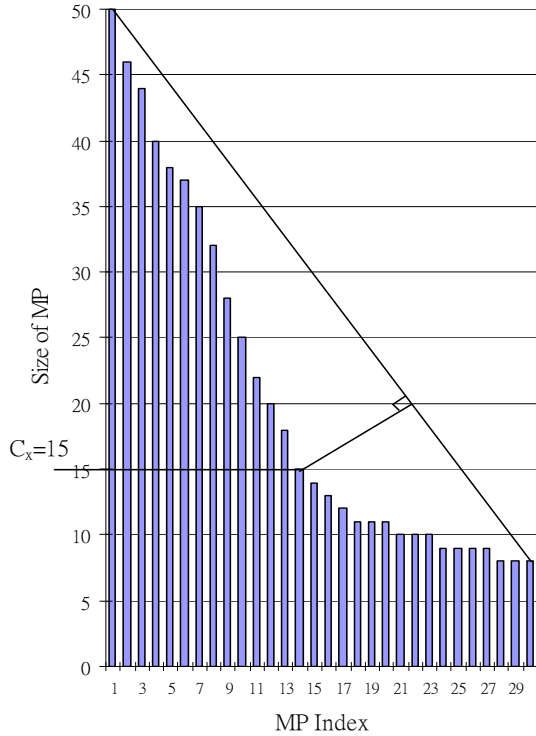


Fig. 2: Calculation of C_x

D. MP Classification

To classify MP, we propose a criterion based on the triangle algorithm [15]. Let N_d denotes the number of observable instances in the d^{th} MP. The criterion is presented by setting a threshold C_x for N_d . For the d^{th} MP, if $N_d > C_x$, then the MP

is classified as an MP-C, otherwise it is classified as an MP-I. Fig. 2 depicts the adaptive calculation of C_x for the results shown in Fig.5. Initially, all MP are ordered in a descending order in terms of N_d . In this case, MP with the size above 15 are classified as MP-C, and the other MP are classified as MP-I.

E. Pedestrian Behavior Prediction

The focus of the proposed method is to predict pedestrian motion behavior in the most appropriate manner based on the MP-C, MP-I and current trajectories, through a multiple prediction hierarchy as depicted in Fig. 3.

Let T_k denotes the observable trajectory of PD_k , and P_m^i and P_n^c represent the m^{th} MP-I and the n^{th} MP-C respectively. T_k is given by $\{T_k^s, T_k^v, T_k^\phi\}$ and P_m^i and P_n^c are given by $\{P_m^{s,i}, P_m^{v,i}, P_m^{\phi,i}\}$ and $\{P_n^{s,c}, P_n^{v,c}, P_n^{\phi,c}\}$, respectively. Let T_k^* denotes the predicted behavior of PD_k in any future motion. T_k^* is also given by $\{T_k^{*s}, T_k^{*v}, T_k^{*\phi}\}$. If T_k is defined up to t , then T_k^* is defined from $t+1$ onward. For illustration convenience, we choose the spatial location feature as an example for presenting the multi-level prediction process. Thus T_k, P_m^i, P_n^c and T_k^* in this case are all simplified into $\{T_k^s, \emptyset, \emptyset\}, \{P_m^{s,i}, \emptyset, \emptyset\}, \{P_n^{s,c}, \emptyset, \emptyset\}$ and $\{T_k^{*s}, \emptyset, \emptyset\}$, respectively. Suppose there are a total of N observable trajectories and M MP with m_2 MP-C and m_1 MP-I ($m_1+m_2=M$).

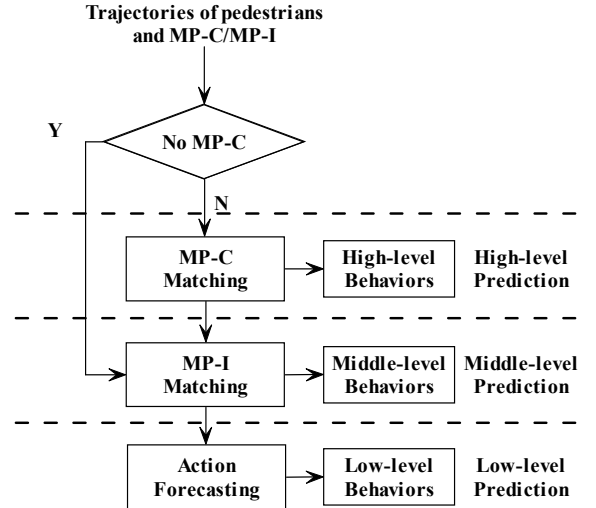


Fig. 3: Block diagram of multi-level prediction

For high level prediction, the trajectories and MP consist of spatial locations of different number of time steps, so the corresponding T_k and P_m^i and P_n^c have different dimensions. Before matching is performed, their dimensions are equalized. To do that, we first segment each P_n^c ($1 \leq n < m_2$) into portions which have the same data dimension with each T_k ($1 \leq k < N$). For example, if T_k has K_k^T time steps, then a portion of P_n^c with K_n^P time steps ($K_n^P > K_k^T$) is denoted by $P_{n(q)}^c$, which is given as:

$$P_{n(q)}^c = t_k^s(q, q + K_k^T), \quad q = 0, 1, \dots, K_n^P - K_k^T. \quad (2)$$

We then use a Gaussian function $f(\mathbf{X}, \boldsymbol{\mu})$ for determining the similarity between each trajectory and each portion of

each MP-C. $f(\mathbf{X}, \boldsymbol{\mu})$ is defined as:

$$f(\mathbf{X}, \boldsymbol{\mu}) = ae^{-\frac{\sum_{i=0}^L (x_i - \mu_i)^2}{2\sigma^2}}, \quad (3)$$

where a is a positive real constant which is set to 1.0, and σ is an error coefficient which is set to 10. Both values are obtained through extensive experimentation. \mathbf{X} and $\boldsymbol{\mu}$ are both vectors, in which x_i and μ_i are their elements, respectively. L is the total number of x_i , which is the same for μ_i after equalization. To perform a match, we simply compute $f(T_k, P_{n(q)}^c)$. T_k is regarded as matched with MP-C P_n^c when it satisfies the following decision function:

$$f(T_k, P_{n(q)}^c) > f(P_{n(q)}^{c'}, P_{n(q)}^c), \quad (4)$$

in which

$$P_{n(q)}^{c'} = \{\mathbf{r}_k[n] + \sigma \mathbf{I}\}, \quad q=0, 1, \dots, K_n^P - K_k^T. \quad (5)$$

where \mathbf{I} is a unit vector and σ is set to the same value as in (3).

When a match is successful, P_n^c is taken as a candidate behavior of T_k . After applying (4) between each trajectory and all MP-C, T_k can be categorized into three cases in terms of the number of candidate behaviors. Let $A(T_k)$ denotes the number of candidate behaviors for T_k . The three cases are:

- Case 1: $A(T_k)=1$. This results in setting the only candidate behavior as the predicted behavior of T_k ;
- Case 2: $A(T_k)>1$. This results in setting the most similar behavior (with the largest $f(T_k, P_{n(q)}^c)$ value) as the predicted behavior of T_k ;
- Case 3: $A(T_k)=0$. The prediction is unsuccessful and T_k needs to be considered in the next level.

If the set of MP-C is *null*, or there exists unpredicted trajectories after the high-level prediction, middle-level prediction is applied based on MP-I. The principle for middle-level prediction is the same as the high-level prediction, whereas the only difference is that P_m^i are used for matching at this level. By replacing P_n^c by P_m^i in the above formulation, we can match the remaining trajectories with MP-I in a similar manner.

If there are still trajectories that cannot be matched with any MP after the middle-level prediction, then a single time step prediction is performed. The next position at time step $s+1$ can be predicted by the following equation:

$$w(s+1) = w(s) + u(s)S + B_s a(s)S^2, \quad (6)$$

where $w(s)$ means the position at time step s , and $u(s)$ and $a(s)$ are corresponding velocity value and acceleration value. B_s is time-dependent and is updated by the adaptive algorithm in [16].

In summary, the algorithmic steps are given below:

- Step 1: If $P_n^c \notin \emptyset$, proceeds to Step 2, otherwise mark all T_k as *first-unpredicted* and go to Step 3.
- Step 2: For each T_k , where $1 \leq k < N$, match it with each P_n^c , where $1 \leq n < m_2$. If match is successful, then output corresponding prediction result T_k^* , otherwise, mark T_k as *first-unpredicted*.

Step 3: For each *first-unpredicted* T_k , where $1 \leq k < n_1$ and $n_1 \leq N$, match it with each P_m^i , where $1 \leq m < m_1$. If match is successful, then output corresponding prediction result T_k^* , otherwise, mark T_k as *second-unpredicted*.

Step 4: For each *second-unpredicted* T_k , where $1 \leq k < n_2$ and $n_2 \leq n_1$, predict a single time step based on an AR model, and output the corresponding prediction result T_k^* .

III. EXPERIMENT

In this section, we present the simulation experiment in a DCE to demonstrate our proposed method. A continuous trajectory of a moving object is generally represented by a series of discrete positions which are recorded at all sampled time steps in the moving process. In this experiment, we assume that the sampling time $T_s=1s$, which is a flexible parameter that can be changed depending on the actual system and environment. The scenario of the simulation experiment is based on the shuttle-bus (the vehicle) for ferrying passengers in an airport terminal, where passengers can freely move around as shown in Fig. 4. We also use spatial location feature as an illustration in this experiment. At each time step, new-born pedestrian trajectories will come into the scene, and also some existing pedestrian trajectories will disappear from the scene.

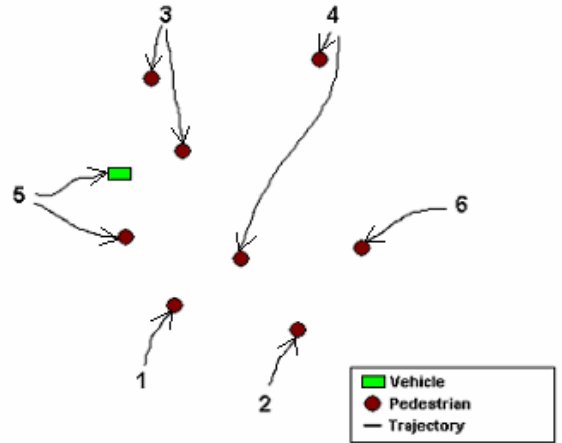


Fig. 4: The scene of the simulation experiment

We track a pedestrian from when it first enters into the scene to when it goes out of the scene and regard its corresponding trajectory in the scene as one training instance for MP clustering. Fig. 5 depicts the observable pedestrian instances and their corresponding clustered MP at $t=200s$.

There are altogether 30 MP which are accordingly clustered. Since there are double-directional trajectories in the scene, two MP are clustered between each pair of entrances in the scene and we use solid-curve and dot-curve for differentiation in Fig. 5. By passing all 30 clustered MP to the MP classification, 13 of them are qualified for as MP-C, as depicted in Fig.2. It should be noted that compared with MP-I, MP-C have better convergent

attributes. Thus, MP-C will be regarded as learned knowledge for high-level prediction which remains in the entire prediction process while MP-I may be upgraded to MP-C later based on more observable pedestrian instances.

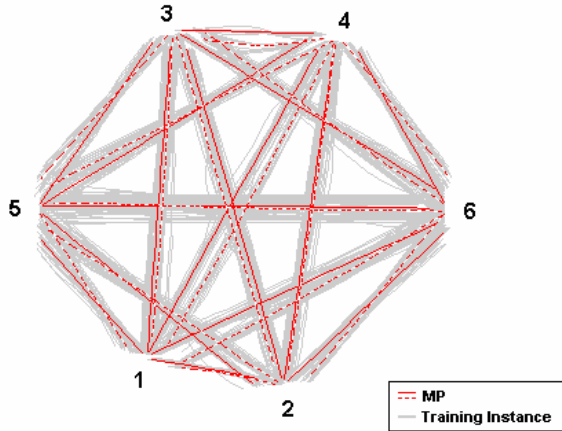


Fig.5: Observable pedestrian instances and clustered MP ($t=200s$)

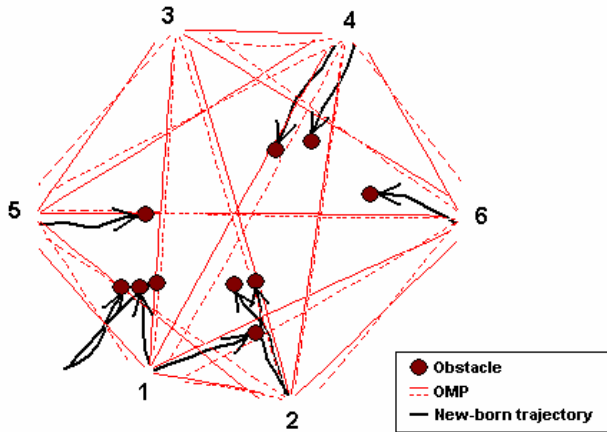


Fig. 6: New-born trajectories for prediction

To further illustrate how prediction is made at different levels, we randomly select the observable trajectories of pedestrians in the scene at $t=220s$, where there are totally 10 observable pedestrian instances (brown-circle) in the scene. Fig. 6 depicts all 10 trajectories (black-curve) and all clustered MP (red-curve). The multi-level prediction results (blue curves) are depicted in Fig. 7. Out of 10 trajectories, 8 can be predicted at the high or middle level, since they match with one of the MP-C or MP-I. It can also be observed that (1) the matching allows slight variations in spatial locations; (2) most of the predictions are long-term; and (3) the essence of the MP determines the predicted behaviors, not the actual spatial locations. For the remaining 2 observable pedestrian instances that could not find a match in either high or middle level, it can be seen that both trajectories are quite different from all existing MP-C/MP-I in both behaviors and spatial locations. In this case, the proposed method makes a single time step action prediction shown by a blue 'x' in Fig. 7.

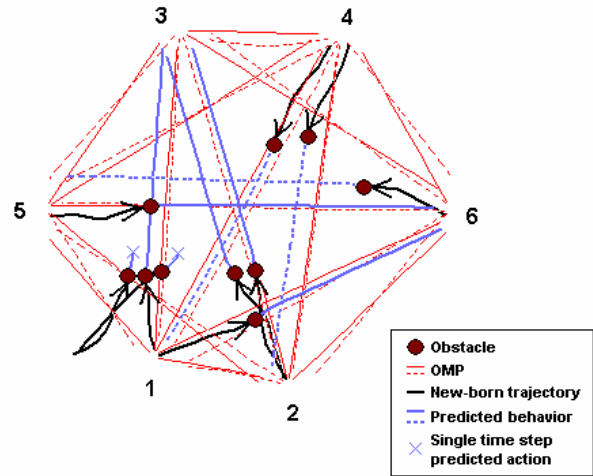


Fig. 7: Multi-level prediction results

In order to test and evaluate the performance of our prediction methods, we have conducted 30 simulation experiments in total 3 different simulation scenarios including the above and compare the predicted behavior with the actual behavior in the simulation. For each predicted trajectory, an error is computed as the ratio between the distance deviation during the whole prediction process which is represented by the accumulated deviated distance between the predicted location and the actual location at the last time step, and the actual total traversed distance. For each simulation experiment, we calculate the average error of all predicted trajectories as the prediction error of the experiment. It is found that the prediction error of each experiment is around 8% and the prediction can be considered as reasonably accurate. The comparison between predicted behavior and actual behavior in the above simulation experiment is shown in Fig. 9.

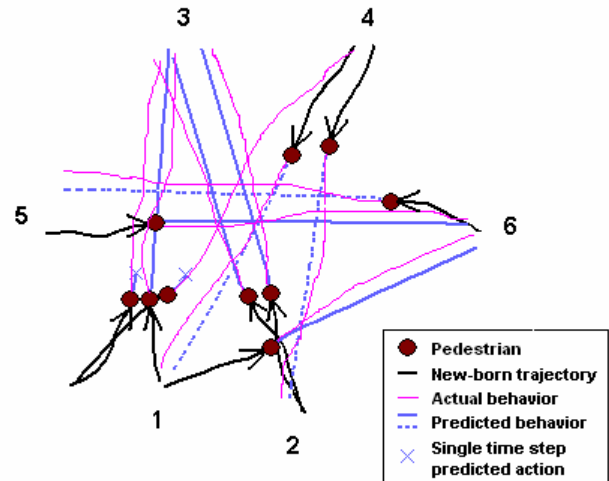


Fig. 8: Comparison between predicted behavior and actual behavior

Fig. 9 shows the application of the prediction results from the proposed multi-level prediction model. For vehicle-to-pedestrian collision avoidance, it is necessary

and important for the vehicle to know where the pedestrians are as well as where they are going to move to in general. With the availability of long-term prediction, the vehicle benefits from being able to ‘look-ahead’ further into the future and make appropriate plan to achieve a global optimal path and to avoid potential collisions and traps.

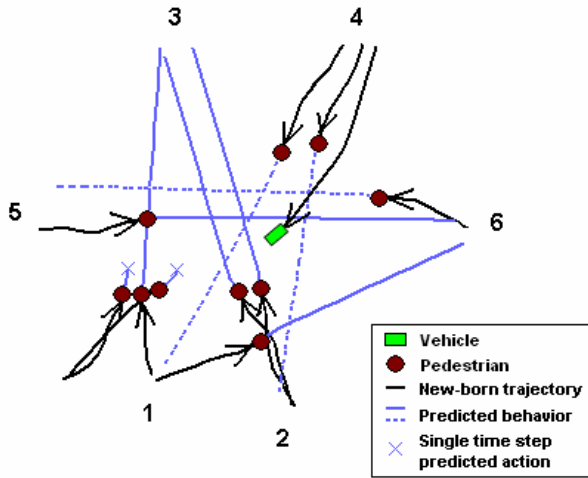


Fig. 9: Application of the proposed multi-level prediction model

IV. CONCLUSION

In this paper, we presented a multi-level prediction model based on MP clustering and classification for long-term pedestrian behavior prediction. From the simulation result, it can be concluded that the proposed method is reasonably effective in deploying high-level, middle-level and low-level predictions for different trajectory behaviors. The main contribution of the proposed method is to offer a viable approach for predicting long-term behavior rather than a single time step prediction, which is more common in existing methods. The ability to make long-term behavior prediction is obviously advantageous, thus, we believe it would have substantial impact to vehicle-to-pedestrian collision avoidance. Furthermore, for some other related application areas, such as crowd control, it should be noted that long-term behavior prediction can also be effective. Based on the general innovative idea and the proposed framework, our future work will mainly focus on three aspects: (1) to conduct extensive simulation study based on velocity and heading angle; (2) to investigate other features to characterize pedestrian behavior; (3) to research online management of MP and to improve the accuracy of pedestrian behavior prediction based on online updated MP.

ACKNOWLEDGEMENT

The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. HKU 7196/06E) and ASM Assembly Automation Ltd.

REFERENCES

- [1] World Health Organization and the World Bank. "World report on road traffic injury prevention: summary". April 2004.
- [2] Christophe F. Wakim, Stephane Capperon, and Jacques Oksman. "Design of Pedestrian Detection Systems for the Prediction of Car-to-Pedestrian Accidents," *IEEE Intelligent Transportation Systems Conference*. Washington, D.C., USA, October 3-5, 2004.
- [3] J. Tani. "Model-based learning for mobile robot navigation from the dynamical systems perspective," *IEEE Transactions on Systems, Man and Cybernetics - Part B*, Vol. 26, No. 3, pp. 421-436, June 1996.
- [4] Q. Zhu. "Hidden Markov model for dynamic pedestrian avoidance of mobile robot navigation," *IEEE Transactions on Robotics and Automation*, Vol. 7, No. 3, June 1997.
- [5] Ashraf Elnagar. "Prediction of moving objects in dynamic environments using Kalman Filters," *Proceedings of IEEE International Symposium on Computational Intelligence in Robotics and Automation*. Banff, Alberta, Canada. July 29-August 1, 2001.
- [6] E. D. Dickmanns, B. Mysliwetz and T. Christians. "An integrated spatio-temporal approach to automatic visual guidance of autonomous vehicles," *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 20, No. 6, pp. 1273-1284, Nov./Dec. 1990.
- [7] A. Chakravarthy and D. Ghose. "Pedestrian avoidance in a dynamic environment: A collision cone approach," *IEEE Transactions on Systems, Man and Cybernetics - Part B*, Vol. 28, No. 5, pp. 562-574, 1998.
- [8] Z. Qu, J. Wang and C. E. Plaisted. "A new analytical solution to mobile robot trajectory generation in the presence of moving pedestrians," *IEEE Transactions on Robotics*, Vol. 20, No. 6, pp. 978-993, December 2004.
- [9] Amalia F. Foka and Panos E. Trahanias. "Predictive control of robot velocity to avoid pedestrians in dynamic environments," *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Las Vegas, Nevada. October 2003.
- [10] Amalia F. Foka and Panos E. Trahanias. "Predictive autonomous robot navigation," *Proceedings of the 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems*. EPFL, Lausanne, Switzerland. October 2002.
- [11] Maren Bennewitz. "Mobile robot navigation in dynamic environments," *PhD thesis*. June 2004.
- [12] D. C. K. Ngai, & N. H. C. Yung, "Double action Q-learning for pedestrian avoidance in a dynamically changing environment", *Proceedings of the 2005 IEEE Intelligent Vehicles Symposium*, Nevada, USA, pp.211-216, ISBN: 0-7803-8962-X, 2005.
- [13] N. H. C. Yung and A. H. S. Lai. "Segmentation of color images based on the gravitational clustering concept," *Optical Engineering*, Vol. 37, No. 3, March 1998.
- [14] Ye Cang. "Behavior-Based Fuzzy Navigation of Mobile Vehicle in Unknown and Dynamically Changing Environment," *PhD thesis*. September 1999.
- [15] Zack, G. W., Rogers, W. E. and Latt, S. A. "Automatic Measurement of Sister Chromatid Exchange Frequency," *Journal of Histochemistry and Cytochemistry*, vol. 25, no. 7, pp. 741-753, 1977.
- [16] M. Shensa. "Recursive least squares lattice algorithms – A geometrical approach," *Automatic Control, IEEE Transactions on*, vol. 26, pp. 695-702, 1981.