Pedestrians' Intention Recognition Method using Hidden Semi-Markov Model: The Case of Crossing the Crosswalk

Junsik Kong, Jaewoong Kang, Jaeung Lee and Mye Sohn
Sungkyunkwan University, Department of Industrial Engineering, Suwon, Korea
jsgong94@skku.edu

***Abstract*— It is very important to ensure that elder people can perform safe outdoor activities, especially crossing the crosswalk. In this paper, we propose a novel system that can recognize intentions of the elder pedestrians in the vicinity of traffic lights to support the safe crossing. In order to recognize the intention, we applied Hidden Semi-Markov Model (HSMM), which is the most widely adopted method in this field of research. Our system consists of three functions: spatial context identification, HSMM-based learning, and intention recognition. To implement our system, we used GPS data collected from sensors embedded in the elder pedestrians' smartphone, traffic lights data collected through Open API, and pre-classified activity data for activity learning. In the experimental section, to evaluate the performance of our system, we conducted experiments to find optimum k of k-prototype clustering and to determine the number of hidden states. The key contribution of this paper is to recognize the intentions from the pedestrians' point of view for the safety of the pedestrians, not the intention of the driver for safe driving of the car.**

***Index Terms*— Human Intention Recognition; Machine Learning; Hidden Semi Markov Model; Elder Pedestrian.**

I. INTRODUCTION

The rapid aging of our society creates a demand to deliver quality daily life for the elderly. To meet the requirement, many researchers have studied on the information communication technologies (ICTs) for the elderly, more specifically the gerontechnologies [1,2]. However, most of the researches related to gerontechnologies are interested to focus on supporting indoor activities, in spite of the elderly over the age 65 in Korea spend nearly 34 percent of their daytime outdoor [3]. As the proportion of outdoor activities by the elderly increases, more and more elderly people are killed or injured while walking [4]. To reduce the risk, we propose a framework to ensure "the ability to cross the road alone" during outdoor activities. The reason why we are interested in "the ability to cross the road alone" is because it is an important indicator of the autonomous capacity of the elderly [5].

To do so, we tried to recognize the intention of the elderly to cross the road, and to adjust the traffic signal of the crosswalk that the elderly want to cross according to the elderly's walking speed. We are particularly interested to understand the intentions of the elderly when crossing the road. Among the research related to human intention recognition (HIR), research that is most relevant to gain intention is one that recognizes the driver's intention. However, the purpose of the research is not on understanding the pedestrians' safety but the prevention of car accidents. Therefore, there is little interest in recognizing the pedestrians' intentions.

To overcome the limitation of existing research, we propose a novel framework that can recognize the pedestrians' intentions, especially in the vicinity of traffic lights to support safe crossing walks, from the pedestrian's point of view. At this time, the pedestrian's intention is recognized using the smartphone sensor carried by the pedestrians. In addition, the data required to control the traffic light were collected using Open API. In order to perform the intention recognition, we adopted Hidden Semi-Markov Model (HSMM), which is most widely adopted method in this field.

The preliminary version of this paper has been presented as a conference paper [6]. As compared with the previous version, the differences and major contributions are as follows.

In the previous version, we performed k-means clustering after data discretization. However, we noticed that the activity data were arbitrary classified categorical data. Therefore, it was difficult to define the distance measure even though it is a discrete variable. Thus, we adopted k-prototype clustering to recognize the activities.

In the previous version, to recognize the intention, we used HMM, in which one observable state determines only one hidden state time-independently. However, in the case of walking, one activity does not determine one hidden state but a series of activities are considered to determine a hidden state. Therefore, a variable corresponding to duration was additionally required. So, we adopted HSMM for recognizing the human's intention.

To verify our framework, we performed the experiments. To do so, we collected real data using application GPSLogger in iPhone X. Through the experiments, we found the optimum k of k-prototype clustering and determined the number of observable states.

The paper is structured as follows. The next section discusses related works. Section 3 offers detailed descriptions about recognition method the pedestrians' intentions in the vicinity of traffic lights to support a safe crossing. The experimentation will be presented in Section 4. Finally, Section 5 puts forth the conclusions and suggests further research.

II. Literature Review

A. Human Intention Recognition

Many researchers have been interested in human intention recognition (HIR) in terms of intelligent systems such as autonomous vehicle. As shown in Table 1, most HIR research had been interested in recognizing the intentions of the drivers or the pedestrians. The mainstream of research shows that the driver grasps the pedestrian intentions. The purpose of previous research is mostly to prevent car accidents. According, this field of research is becoming more active due to the proliferation of autonomous vehicles.

Table 1
Classification of HIR Research

Perspective / Recognition target	Driver	Infrastructure (Camera on the traffic light)	Pedestrian
Drivers' intention	[8], [9], [10], [12], [14]	[19], [20]	-
Pedestrians' intention	[7], [15], [16], [17], [18]	[21]	This research

However, another related area of research is the pedestrians' jaywalking caused by improper traffic light control as well as cars, which are the causes of the elderly pedestrians' accidents on the crosswalk. Therefore, to address this issue, it is very important to recognize the pedestrian's intention from the viewpoint of the pedestrian. Unfortunately, studies in this field of research are still lacking, implying that research to ensure the pedestrians' safety walking from the pedestrians' point of view is unique and novel.

B. HIR Methods

To recognize the humans' intention, the researchers applied a variety of classification methods. As summarized in Table 2, each method has its advantages and disadvantages.

Table 2
Comparison of HIR Method

Method	Characteristics	Ref
Bayesian Network (BN)	– Pros: Strength in small size and incomplete data; Fast response – Cons: No feedback loop due to acyclic property of BN; Difficulty of collection and structuring the expert knowledge	[7], [22]
Support Vector Machine (SVM)	– Pros: High Accuracy; Low cost; Robust from data – Cons: Low efficiency at high dimension data; Need to preprocess (e.g., feature selection); Inadequate to time-dependent data	[12], [13], [14], [19], [20]
Neural Network (NN)	– Pros: Black box; Unable to understanding important features; Need to large data; High computational cost – Cons: No feedback loop due to acyclic property of the BN; Difficulty of collection and structuring the expert knowledge	[11], [18], [19]
Hidden Markov Model (HMM)	– Pros: Very useful for sequential data analysis – Cons: No guarantee the convergence to the global minimum; Unable to consider the duration of a state	[8], [9], [10], [19], [23]

The pedestrians' intention to be recognized in this research has a sequential characteristic [23]. However, SVM and NN are inadequate for this research because they assume that intentions are time independent. To solve this problem, He L. is applied the Hidden Markov Model (HMM), which has the strength in dynamic time series modeling [10]. However, in HMM, one observable state determines only one hidden state. However, in the case of walking, one activity does not determine one hidden state, but a series of activities are gathered to determine one hidden state. Therefore, a variable corresponding to duration is required. In this case, the Markov model with duration is called the semi-Markov model. For the purpose of this study, we adopted HSMM for recognizing human intention.

III. Overall Architecture

A. System Architecture

The architecture of this system consists of two phases, such as learning and recognition. Learning phase carries out learning about activity patterns of the elderly pedestrian around the Identification Module (SCIM), Ordered Context-augmented Activity Clustering Module (OCAM), and HSMM Model Generation Module (HMGM). In the recognition phase, our system tries to recognize the pedestrian's intention using real-time data. The overall system architecture is depicted in Figure 1.

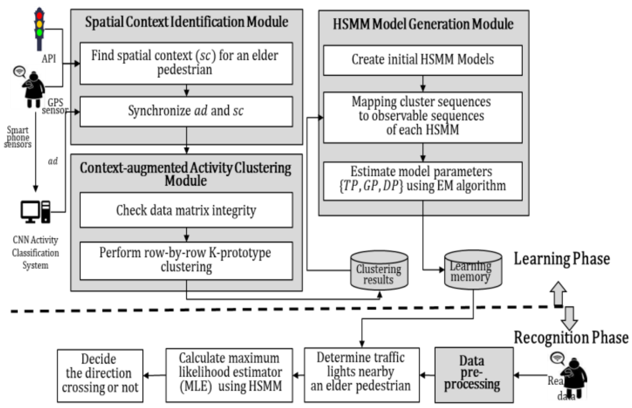

Figure 1: Overall System Architecture

B. Learning Phase

1) Identification of Spatial Context

In order to recognize the crossing intention of the pedestrians, it is essential to clearly define the current context, in which the pedestrian is exposed. It is especially important to clearly identify the spatial context required for the understanding of the relationship between a pedestrian and traffic lights. Prior to the discussion, the variables needed to define the spatial context are depicted in Figure 2.

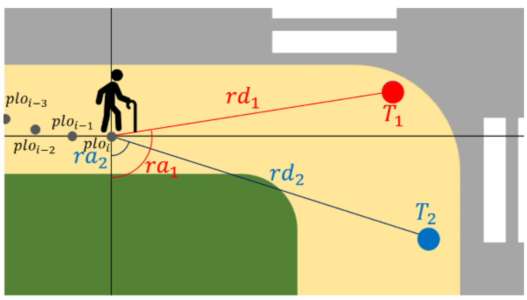

Figure 2: Description of the Variables

pl_i: i^{th}location coordinates of the pedestrian. If $\mathrm{i} = \mathrm{c}$, it means the current location of the pedestrian (pl_c). It is represented as $pl_i = (lo_i, la_i)$.

T_j : j^{th} traffic light coordinates near the pedestrian ($\mathrm{j} = 1,2$). It is represented as $\mathrm{T}_j = (lo_j, la_j)$.

rd_{ij}: Euclidean distance between plo_i and T_j ($\mathrm{i} \geq 1, \mathrm{j} = 1,2$).

ra_{ij}: Relative angle between plo_i and T_j($\mathrm{i} \geq 1, \mathrm{j} = 1,2$).

pc_{ij}: Polar coordinates for p_i and T_j. $\mathrm{pc}_{ij} = (\mathrm{rd}_{ij}, \mathrm{ra}_{ij})$.

st_i: Timestamp of i^{th}spatial context

At this time, the origin of polar coordinates is equal to the current location of the pedestrian. The polar coordinates are derived as follow (Figure 3).

$\overrightarrow{\mathrm{WD}}$: vector of walking direction
$\overrightarrow{\mathrm{PA}}$: vector of polar axis
$\overrightarrow{\mathrm{TV_j}}$: direction vector of traffic lights
p: pole, $p = pl_c$
δ: threshold of rd_{ij}

Begin Polar Coordinates Calculation Process

While $((rd_{i1} < \delta)\ and\ (rd_{i2} < \delta))${
  $\mathrm{p} \leftarrow pl_c$;
  $\overrightarrow{\mathrm{WD}} \leftarrow [pl_{c-1}\ p]$
  $\overrightarrow{\mathrm{PA}} \leftarrow \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \times \overrightarrow{\mathrm{WD}}$;; rotate
  For ($j = 1\ ; j \leq 2\ ; j++$)
    $\overrightarrow{\mathrm{TV_j}} \leftarrow [T_j - \mathrm{p}]$
    $ra_{cj} \leftarrow \cos^{-1}\frac{\overrightarrow{\mathrm{PA}} \cdot \overrightarrow{\mathrm{TV_j}}}{|\overrightarrow{\mathrm{PA}}| \times rd_{\mathrm{cj}}}$
    $pc_{cj} \xleftarrow{append} (rd_{cj}, ra_{cj})$
  EndFor}

Return $pc_{ij}\ for\ all\ i\ andj$
End Process

Figure 3: Derivation Algorithm of Polar Coordinates

Based on the above notations and the algorithm, the i^{th} spatial context is defined as follow.

Definition 1 i^{th} Spatial Context (sc_i) is represented as follows.

$$sc_i = \{pl_i, pc_{i1}, pc_{i2}, st_i\} \tag{1}$$

However, it may be different from the time span of the derived polar coordinates and the activity data received from the activity recognition module. Therefore, in order to use two kinds of data having different time span as input values to the HMM, it is necessary to synchronize them into the same time-window. The received activity data from the action data recognition module are defined as follow.

Definition 2 k^{th} activity data (ad_k) is represented as follows.

$$ad_k = \{ac_k, at_k\} \tag{2}$$

where ac_k is k^{th}activity of the pedestrian ($ac_k \in$ {walking, standing, turning left, turning right}). at_k is time stamp of k^{th} activity.

Definition 3 Context-augmented activity data matrix is represented as follow.

$$M_l = \begin{bmatrix} pl_1\ pc_{11}\ pc_{12}\ st_1\ ac_1 \\ \vdots \\ pl_l\ pc_{l1}\ pc_{l2}\ st_l\ ac_l \end{bmatrix} \tag{3}$$

Based on the above definition, the synchronization procedure of the time span is summarized in Figure 4.

temp= ∅

Begin Data Synchronization Process

For ($k = 1; k \geq 1; k++$)
  If $at_k > st_{c-1}$
    Then $m \leftarrow k$
      Break
  End If
End For
For ($k = m; k \geq 1; k++$)
  If $at_k > st_c$
    Then $n \leftarrow k$
      Break
  End If
End For
For ($k = m; k < n; k++$)
  temp $\xleftarrow{append} ac_k$
EndFor
$sc_{\mathrm{c}} \xleftarrow{append}$ mode(temp)
$M_{\mathrm{c}} \xleftarrow{append} \mathrm{sc_c}$

Return M_c
End Process

Figure 4: Synchronization procedure of time spans

2) Ordered Context-augmented Activity Clustering

In order to recognize the spatial context, we collected the GPS sensor data embedded in the smartphone of the elder pedestrians, and traffic lights-related data collected using open API. At this time, due to the device errors and sensor degradation, the collected sensor data are incomplete. Furthermore, the collection period of API (greater than 1sec) differs from the collection period of sensor data of the smartphone (1/40sec). Thus, in order to generate data matrix for clustering, data refinement is required. If the sensor data value is missing, it may be replaced with the previous data or the subsequent data value, but the data collected by the API requires careful approach. To treat the missing data collected through open API, data interpolation should be performed. Data integrity is solved by processing null data.

After confirming the data integrity, we performed the clustering to recognize the activities. At this time, context-augmented matrix contains not only discrete domains (e.g., activity) but also continuous domains (e.g., distance and angle). Generally, clustering in the continuous domain uses the distance function as a cost function. It is called k-means clustering. However, activity data is the arbitrary classified categorical data. Therefore, it is difficult to define the distance measure even though it is a discrete variable. To complement the distance measure of the categorical data, k-modes clustering using dissimilarity as the distance has appeared [25]. Recently, k-prototype clustering, which integrates k-means clustering and k-modes clustering has emerged [26]. The context-augmented activity data matrix consists of p-dimensional context data as numerical data and 1-dimensional activity data as categorical data. In this light, the k-prototype clustering is appropriate for our data.

In the k-prototype clustering, the distance function is defined as follows.

$$d(\mathrm{X},\mathrm{Y}) = \sum_{j=1}^{p}\left(x_j - y_j\right)^2 + \gamma \times \delta(x_{p+1}, y_{p+1})$$
$$(x_j, y_j) = \begin{cases} 0 \ (x_j = y_j) \\ 1 \ (x_j \neq y_j) \end{cases} \tag{4}$$

where $x_j, y_j (1 \leq j \leq p)$ are the numerical data, x_{p+1}, y_{p+1} are the categorical data. δ is the dissimilarity function of two categorical data. γ is the hyper parameter to avoid favoring either the type of attribute.

The disadvantage of k-prototype clustering is that there is no way to obtain an optimum k, unlike k-means clustering, because γ must be adjusted. Therefore, the optimum k is obtained experimentally so that the recognition accuracy could be high.

3) HSMM-based Learning

HMM is a probabilistic model, which is used for recognizing the hidden states from the observable sensor data [18]. When the HMM is used in the intention recognition, the intentions are represented as the hidden states and the sensor data is used to depict the observable states. In HMM, an intention is determined by the transition probability which is time-independent. However, in the case of crossing the road, the intention should be determined considering the duration of a prior intention. For instance, whether the pedestrian would cross the road depend on how long the pedestrian has kept the intention of waiting around the traffic light.

Therefore, a variable corresponding to duration is required. The Markov model with duration is called semi-Markov model. This model is transformed into HSMM in the human intention recognition.

In this research, three different types of HSMM models are learnt since the direction of elder pedestrian at crossroads can be forward, rightward and leftward. For each model, the intentions are defined as a sequence of the hidden states like approaching to traffic signal, waiting, crossing the street, and moving away. For example, an HSMM of crossing the street is depicted in Figure 5.

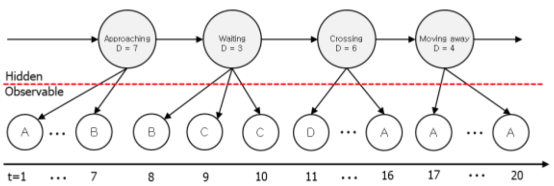

Figure 5: Observation sequence of an intention named crossing the street

In order to use the Context-augmented activity data as the observable states, this data should be transformed to discrete type. Discretization is performed by k-prototype clustering of the sensor data. The k clusters are used to train three HSMMs with differently labeled hidden states. At this time, the three kinds of intentions are crossing the right crosswalk, crossing the left crosswalk, and crossing the front crosswalk.

To recognize the pedestrians' intention, the HSMM is represented by three matrices, $\lambda = (\mathrm{A}, \mathrm{B}, \pi)$. The matrix $\mathrm{A} = a_{(i,d'),(j,d)}$ represents the transition probability that the hidden state will change from i_i to i_j. The matrix $\mathrm{B} = b_{i,d}(o_{t+1:t+d})$ specifies the probability that the system will generate the observable states $o_{t+1:t+d}$ with hidden state i_i for duration d. And $\pi_{i,d}$ is the distribution of probability that any given state is the initial state of the hidden semi Markov process. A, B, and π are simply represented as follow.

$$\begin{aligned} a_{(i,d'),(j,d)} &= p\left(H_{[t+1:t+d]} = i_j \middle| H_{[t-d'+1:t]} = i_i\right) \\ b_{i,d}(o_{t+1:t+d}) &= p\left(O_{t+1:t+d} \middle| H_{[t+1:t+d]} = i_i\right) \\ \pi_{i,d} &= p(H_{[t-d+1:t]} = i_i) \end{aligned} \tag{5}$$

Given $\lambda = (\mathrm{A}, \mathrm{B}, \pi)$, the probability that the intention H is recognized is as follow. At this time, the intention is one of two cases like crossing the street or passing the crosswalk.

$$p(H|\lambda) = \sum_{for\ all\ O} p(H|\mathrm{O}, \lambda)p(O|\lambda) \tag{6}$$

where the set $H = \{(i_1, d_1), (i_2, d_2), (i_3, d_3), (i_4, d_4)\}$ is the expression of the hidden states and observable states are expressed as $O = \{o_{1:j}, \ldots o_{m:n}\}$.

C. Recognition Phase

In the recognition phase, it determines whether the pedestrian has the intention to cross the street by comparing real-time sensor data related to walking with HSMM learning results. To do so, the recognition phase performs three tasks sequentially. First, it calculates the distance between the pedestrian and the traffic lights. At this time, the location of the pedestrian is periodically identified using the GPS sensor embedded in the elder pedestrian smartphone. When the distance is below the threshold (δ), it starts collecting the walking data of the pedestrian in real time. Using the sensor data, it creates a spatial context by predefined time span. Second, it performs preprocessing on the sensor data in order to synchronize the time spans of the activity data received from the activity recognition module and data collection time. After the two tasks are finished, we can identify the traffic lights in two directions that are related to the pedestrian's intention. As a next step, we calculate Maximum Likelihood Estimator (MLE) for each direction using real-time data. Finally, our system decides whether the elder pedestrian will cross the crosswalk in a certain direction or just pass by.

IV. PERFORMANCE EVALUATION

To show the application example of our framework, we conducted the simulation. To collect the data, the experiment was conducted on the roads around Sungkyunkwan University. At the viewpoint of the subject, the road is curved to the left and the crosswalks are located at the front and the right. Thus, in this case, models for crossing the forward crosswalk, crossing the right crosswalk and passing by would be the proper combination. The subject walked in three directions naturally as usual. Each observation time is at least 1 minute to maximum 3 minutes.

Collection frequency of GPS data was set to 1Hz and an observation starts after the GPS data is stabled. The number of observations is the same on each direction. The application GPSLogger was used on the smart phone (iPhone X). Data pre-processing was performed in Python and the Models were generated in R.

In order to create a HSMM, we need to decide the number of the hidden states. Since there is no fancy theory and the number of states highly depends on the domain, the number should be decided in the experimental way. To decide it, the log-likelihood values of cluster sequences are calculated changing the number of the hidden states. Then, the number with the most log-likelihood value on each model would be selected. The result is depicted in Figure 6.

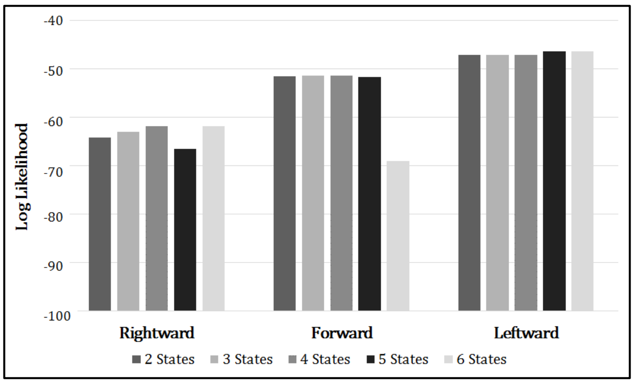

Figure 6: Log Likelihood values for each model

According to Figure 6, four hidden states would be appropriate.

As mentioned above, since the number of clusters in k-Prototype method also has no theoretical solution, the number should be decided in the experimental way too. The following tables show how the result of the recognition is verified according to the number of clusters (k).

Table 3:
Log-Likelihood Values When k=8

Model / Data	Log-Likelihood Leftward	Forward	Rightward
Leftward	-24.7584	-44.0192	-25.5406
Forward	-56.3914	-54.5724	-58.4552
Rightward	-49.2543	-65.4980	-24.7117

Table 4:
Log-Likelihood Values When k=12

Model / Data	Log-Likelihood Leftward	Forward	Rightward
Leftward	-47.1364	-58.6592	-70.4692
Forward	-53.9968	-51.4945	-62.0207
Rightward	-76.3990	-101.5220	-61.8045

As shown in Table 3, k = 8 is not proper for our dataset. The system predicts the pedestrian would cross the forward crosswalk although the pedestrian actually passed to the left. Continuously changing the number k and calculating the log-likelihood values, we could find the optimum k to be 12. Under the condition of 4 hidden states and 12 kinds of clusters, the system could successfully predict the pedestrian's intention.

V. Conclusion and Further Research

In this paper, we propose a novel system that can recognize intentions of the elder pedestrians in the vicinity of the traffic lights to support a safe crossing. The architecture of this system consists of two phases such as learning and recognition. Learning phase carries out the learning about activity patterns of the elder pedestrian around the crosswalk. To do so, it composes three modules: Spatial Context Identification Module (SCIM), Ordered Context-augmented Activity Clustering Module (OCAM), and HSMM Model Generation Module (HMGM). In the recognition phase, our system tries to recognize the pedestrian's intension using real-time data.

Contributions of our research can be summarized as follows. First, we tried to recognize the pedestrians' intentions from the pedestrian's point of view. Second, to recognize the elder pedestrians' intention around the crosswalk, we used only the sensor data embedded in the smartphone carried by the elder pedestrians without the help of guardians or additional devices. Last but not least, we used HSMM to model the activity to recognize the intention.

For further research, we will try to integrate our system to LOD to find the paired information of traffic lights. In addition, we can try to combine this study with Heinz's intention recognition level [27].

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF-2016 R1D1A1B03932110).

References

[1] Geman, Oana, et al. "Challenges and trends in Ambient Assisted Living and intelligent tools for disabled and elderly people." Computational Intelligence for Multimedia Understanding (IWCIM), 2015 International Workshop on. IEEE, 2015.
[2] Kupiainen, Tiina, and Tiina Jansson. "Aged People's Experiences of Gerontechnology Used at Home: A narrative literature review." (2017).
[3] Statistics Korea [KOSTAT], 2015 Statistics on the Aged, 2015, Retrieved at http://kostat.go.kr/
[4] Korea Traffic Accident Analysis System [KoROAD], Traffic accident statistics summary for 2015, Retrieved on July 27, 2015
[5] Asher, Laura, et al. "Most older pedestrians are unable to cross the road in time: a cross-sectional study." Age and ageing 41.5 (2012): 690-694.
[6] Kong J., et al., "Recognition of pedestrians' intention around traffic lights using Hidden Markov Model", In International Conference on Internet (ICONI) 2017, 2017, December.
[7] Hashimoto, Yoriyoshi, et al. "Probability estimation for pedestrian crossing intention at signalized crosswalks." Vehicular Electronics and Safety (ICVES), 2015 IEEE International Conference on. IEEE, 2015.
[8] Tran, Duy, et al. "A Hidden Markov Model based driver intention prediction system." Cyber Technology in Automation, Control, and Intelligent Systems (CYBER), 2015 IEEE International Conference on. IEEE, 2015.
[9] Berndt, Holger, Jorg Emmert, and Klaus Dietmayer. "Continuous driver intention recognition with hidden markov models." Intelligent Transportation Systems, 2008. ITSC 2008. 11th International IEEE Conference on. IEEE, 2008.
[10] He, Lei, Chang-fu Zong, and Chang Wang. "Driving intention recognition and behaviour prediction based on a double-layer hidden Markov model." Journal of Zhejiang University SCIENCE C 13.3 (2012): 208-217.

[11] Völz, Benjamin, et al. "A data-driven approach for pedestrian intention estimation." Intelligent Transportation Systems (ITSC), 2016 IEEE 19th International Conference on. IEEE, 2016.
[12] Huang, Xianyi. "Driver lane change intention recognition by using entropy-based fusion techniques and support vector machine learning strategy." (2012).
[13] Köhler, Sebastian, et al. "Early detection of the pedestrian's intention to cross the street." Intelligent Transportation Systems (ITSC), 2012 15th International IEEE Conference on. IEEE, 2012.
[14] Kumar, Puneet, et al. "Learning-based approach for online lane change intention prediction." Intelligent Vehicles Symposium (IV), 2013 IEEE. IEEE, 2013.
[15] Schulz, Andreas Th, and Rainer Stiefelhagen. "Pedestrian intention recognition using latent-dynamic conditional random fields." Intelligent Vehicles Symposium (IV), 2015 IEEE. IEEE, 2015.
[16] Diederichs, Frederik, et al. "Application of a Driver Intention Recognition Algorithm on a Pedestrian Intention Recognition and Collision Avoidance System." UR: BAN Human Factors in Traffic. Springer Vieweg, Wiesbaden, 2018. 267-284.
[17] Schulz, Andreas. "Video-based Pedestrian Intention Recognition and Path Prediction for Advanced Driver Assistance Systems." (2016).
[18] Dominguez-Sanchez, Alex, Miguel Cazorla, and Sergio Orts-Escolano. "Pedestrian movement direction recognition using convolutional neural networks." IEEE Transactions on Intelligent Transportation Systems 18.12 (2017): 3540-3548.
[19] Zhao M., Käthner D., & Jipp M., “Modeling driver intention and behavior at roundabouts”, Interdisziplinärer Workshop Kognitive Systeme 2015 (4), 2015, March.
[20] Salomonson, Ivar, and Karthik Murali Madhavan Rathai. Mixed driver intention estimation and path prediction using vehicle motion and road structure information. Diss. Master thesis, Chalmers University of Technology, Gothenburg, Sweden, 2015.
[21] Uusitalo, Laura. "Advantages and challenges of Bayesian networks in environmental modelling." Ecological modelling 203.3-4 (2007): 312-318.
[22] Krauthausen, Peter, and Uwe D. Hanebeck. "Intention recognition for partial-order plans using dynamic bayesian networks." Information Fusion, 2009. FUSION'09. 12th International Conference on. IEEE, 2009.
[23] Rasouli, Amir, Iuliia Kotseruba, and John K. Tsotsos. "Agreeing to cross: How drivers and pedestrians communicate." Intelligent Vehicles Symposium (IV), 2017 IEEE. IEEE, 2017.
[24] Quintero, Raúl, et al. "Pedestrian intention recognition by means of a Hidden Markov Model and body language." Intelligent Transportation Systems (ITSC), 2017 IEEE 20th International Conference on. IEEE, 2017.
[25] Chaturvedi, Anil, Paul E. Green, and J. Douglas Caroll. "K-modes clustering." Journal of Classification 18.1 (2001): 35-55.
[26] Huang, Zhexue. "Extensions to the k-means algorithm for clustering large data sets with categorical values." Data mining and knowledge discovery 2.3 (1998): 283-304.
[27] Heinze, Clint. Modelling Intention Recognition For Intelligent Agent Systems. No. Dsto-Rr-0286. Defence Science And Technology Organisation Salisbury (Australia) Systems Sciences Lab, 2004.