HAL
open science

Natural Vision Based Method for Predicting Pedestrian Behaviour in Urban Environments

Pavan Vasishta, Dominique Vaufreydaz, Anne Spalanzani

To cite this version:

Pavan Vasishta, Dominique Vaufreydaz, Anne Spalanzani. Natural Vision Based Method for Predicting Pedestrian Behaviour in Urban Environments. IEEE 20th International Conference on Intelligent Transportation Systems, Oct 2017, Yokohama, Japan. 10.1109/ITSC.2017.8317848 . hal-01561029

HAL Id: hal-01561029

https://inria.hal.science/hal-01561029

Submitted on 25 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Natural Vision Based Method for Predicting Pedestrian Behaviour in Urban Environments

Pavan Vasishta[1], Dominique Vaufreydaz[2] and Anne Spalanzani[3]§

Abstract

This paper proposes to model pedestrian behaviour in urban scenes by combining the principles of urban planning and the sociological concept of *Natural Vision*. This model assumes that the environment perceived by pedestrians is composed of multiple potential fields that influence their behaviour. These fields are derived from static scene elements like side-walks, cross-walks, buildings, shops entrances and dynamic obstacles like cars and buses for instance. Using this model, autonomous cars increase their level of situational awareness in the local urban space, with the ability to infer probable pedestrian paths in the scene to predict, for example, legal and illegal crossings.

1 INTRODUCTION

As the race to attain and deploy fully autonomous vehicles on urban roads heats up, the concept of Situational Awareness (SA) takes centre stage. Situational awareness is the natural human ability to understand, react and predict the environment based on previously learnt parameters, the utilisation of which is most frequently seen while driving. Human drivers need to balance different variables – speed, route selection, positions of pedestrians, cyclists, other cars, etc. – while trying to predict their future states. In fact, errors in maintaining situation awareness are the most frequent cause of errors in real-time tasks such as driving [1] and can be attributed to many accidents.

Situational Awareness can be described in three incremental abstract levels [2] as *Perception*, *Comprehension* and *Projection*. A human driver in an urban street cycles through these levels continuously. Objects and elements– pedestrians, obstacles, other cars, cross-walks, interesting areas– in the environment are identified. These elements are contextually understood with regard to the environment that they are in– the answer to the question "Why is that element there?" For example, the answer to "Why is there a cross-walk on the street" is to facilitate a crossing from one side of the road to another. Finally, the objects and elements are understood together and their future interactions are projected with a certain probability: a cross-walk *may* be used by a pedestrian if he/she is close to it. A human driver's specific course of action is decided by these continuously evolving projections.

The main motivation of this work is to increase the situational awareness of an autonomous car in the context of driving in urban streets. A major driving force is the adoption of sociological ideas for understanding pedestrian behaviour in inner city areas. Pedestrian behaviour has been postulated to be a function of the built environment; i.e. their movement is a consequence of the presence of certain positive and negative *attractors* [3]. This behaviour, called Natural Motion, is an extension of Gibson's *Natural Vision* which envisages human behaviour as wanting to move in a direction that interests them the most in their field of view [4]. These positive attractors, here called "Points of Interest (POI)", may be present as an element in the scene. They can be monuments, places of public interest, public transportation... Other, more common, POIs are areas of commercial interest - stores, restaurants, etc., that are seen very frequently in an urban centre [5]. The presence of these POIs in any scene influences the behaviour of pedestrians within it. Understanding these influences allows the autonomous car to perform actions that are instinctive in a human driver - project pedestrian future states and intuit areas of legal and illegal crossings.

*1 Pavan Vasishta is a PhD student in the CHROMA team (Univ. Grenoble Alpes, Inria, 38000 Grenoble, France, email: Pavan.Vasishta@inria.fr)

†2 Dominique Vaufreydaz is an associate professor in the Pervasive Interaction team (Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP*, LIG, 38000 Grenoble France, email: Dominique.Vaufreydaz@inria.fr)

‡3 Anne Spalanzani is an associate professor in the CHROMA team (Univ. Grenoble Alpes, Inria, 38000 Grenoble France, email: Anne.Spalanzani@inria.fr)

§* Institute of Engineering Univ. Grenoble Alpes

The major contribution of this work is the creation of a new framework for quick comprehension of urban streets. It also forms a base to project future states of pedestrians without the need of their presence in the scene, analogous to a human driver's intuition. We model the scene as attractive and repulsive potential fields. The novelty of our approach is the introduction of POIs whose attractiveness influences pedestrian behaviour.

The paper is divided into five sections. Section II deals with related work in the field of pedestrian motion prediction, followed by section III, the theoretical basis of the framework and the methods used to project future states. Section 4 discusses the implementation of this framework, its results and validation of a conducted experiment. Section 5 concludes this paper with a discussion on the current work and envisaged future work.

2 RELATED WORK

As far as the authors know, there has been little to no work done in the field in accounting for Point of Interest influences in urban pedestrian prediction. Much work, however, has been done in modelling and prediction of pedestrian routes and route-choice behaviour [6]. Most pedestrian behaviour prediction algorithms depend on learning frameworks based off of observed data. Modelling the inherent pedestrian variables as one technique. A data driven approach on this, minimising an energy function that accounts for many personal factors like speed, grouping etc., can be found in [7]. A similar approach, based on the *Social Force Model* can be seen in [8].

Others use Markov Decision Processes (MDPs) and its variants to predict the beliefs of pedestrian crossings in a scene [9]. Destinations are assumed to be known. In close spirit to our work are [10], [11] and [12]. These works build a cost function based on the environment. [10] learns the cost function of the environment via observed trajectories. An MDP is solved with rewards based on observed trajectories as well with known destinations. Working on static scenes, it requires previously observed trajectories to model the cost function and thus is infeasible for rapidly changing scenes on autonomous vehicles. [11] continues this work by semantically segmenting the observed environment to construct a cost function fed to an hMDP to predict pedestrian positions, even without pedestrian observations. While this knowledge is transferable, it is computationally expensive. The computation problems are solved in [12], yet being applied only for static scenes with learned trajectories. Last, [13] looks at the distance of the pedestrian from the kerb, position and velocity of the car from the cross-walk to learn an "*Inner city model*" to predict pedestrian crossings.

These works require learning from observed data to predict pedestrian positions. Our framework envisages to solve the issues of computational complexity and dynamicity while considering environmental and social factors. It also does *not* use learning for building environmental models which makes it very useful to deploy in unknown environments, with very few dependencies.

3 THEORETICAL FRAMEWORK

Pedestrian crossing behaviour in urban areas can be classified into two broad categories - legal crossings and illegal crossings. Legal crossings are such movements of a pedestrian that account for the safest path from one side of the street to the other. These generally happen on a cross-walk. An illegal crossing is an abnormal behaviour wherein the pedestrian decides to not take the cross-walk to cross the street.

In a structured urban environment, for legal crossings to occur, certain assumptions are made:

- The edges of the road repel pedestrians such that their paths are restricted to the side-walk.
- A cross-walk acts as a conduit between the two sides of the street and offers no resistance to crossing
- The road acts as a barrier for crossing, repelling pedestrians towards the side-walks.
- Static and Dynamic obstacles on the road are repulsive in nature, increasing the resistance of the road and pushing back pedestrians towards side-walks.
- Side-walks offer no resistance to pedestrian movement.
- Points of Interest are a reason for pedestrians to cross from one side of the street to another.

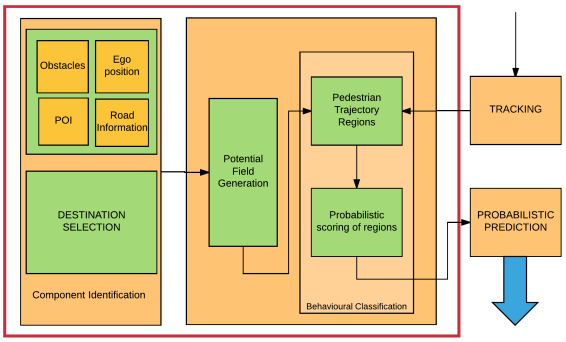

Figure 1: Architecture of the Framework. Only the red block has been implemented in this work

An illegal crossing occurs when at least one of these assumptions is violated. Predicting these areas of illegal crossings leads to a higher level of SA. Looking at these assumptions, it can be seen that a system of potential fields [14] can be a good fit as a model for explaining urban behaviour. Each of the assumptions made earlier can be represented as a function of a potential field.

The architecture of the proposed framework can be found in Fig. 1. From the observed scene, certain informations need to be extracted. These informations are road width, number of lanes, the closest Points of Interest (POI) from the observer, their orientation, the position of the closest cross-walk, etc. Static and Dynamic obstacles on the road also need to be identified.

Destination points in the scene are chosen and fed forward along with the scene information to the Potential field generator. This potential field generation step generates a grid with a potential "Map" based on the extracted data. The generated "Map" is used by the Behaviour classification module, first to generate pedestrian trajectory regions and to score these regions with a probability.

Firstly, to demonstrate the model, it is assumed that the scene under consideration is well-structured. This implies that there is an observable demarcation between the road and the side-walk, the road and the cross-walk and that the lanes on the road are easily observable. It is also assumed that the width of the road (L_{road}), the width of the lane (L_{Lane}), and the POI positions are known, can be computed using sensors embedded into the autonomous car or can be retrieved from a global map. Given the position of the ego vehicle, the distance to the cross-walk and the side-walk can be extrapolated from known information. Considering the pedestrian as a self-driven particle under the influence of attractive and repulsive forces, a potential field can be constructed which produces certain motion behaviours [5].

A destination, by definition, draws a self driven particle towards it. Thus, a POI can be a destination. Conversely, all destinations in a scene are points of interests. Making this assumption, the viable ends of the observed scene are designated as POI and the potential field is recalculated.

A grid is defined for the observed area with its origin at the top left corner and extending to (X_m, Y_m), the maximum grid values with a specified grid resolution. Each cell on the grid can take the attributes *road*, *cross-walk*, *empty*, *POI*, *obstacle* and *edge* and a range of values between 0 and 1. Taking inspiration from work done in [15], the resultant model is a linear function of all component potential values at each cell in the grid. Thus,

$$\begin{aligned}\mathbf{U_{total}} = \mathbf{U_{Edge}} + \mathbf{U_{Road}} + \mathbf{U_{Obs}} + \\ \mathbf{U_{CW}} + \mathbf{U_{POI}}\end{aligned} \tag{1}$$

where $\mathbf{U_{Edge}}, \mathbf{U_{Road}}, \mathbf{U_{Obs}}, \mathbf{U_{CW}}$ and $\mathbf{U_{POI}}$ are potentials associated with the road edges, the road, obstacles on the road, the cross-walk and the POIs.

3.1 Computing Potentials

For each cell in the grid, the different potentials can be calculated as follows.

3.1.1 Edge Potential

The edge potential must repel pedestrians towards the center of the side-walk. An illegal crossing occurs when the self driven particle can exert enough force to overcome this potential. For each cell with a center (x, y), the value of the potential is defined by

$$\mathbf{U^{ij}_{Edge}} = \frac{1}{2}\eta\left(\frac{1}{\rho(x^{ij}, y^{ij})}\right) \tag{2}$$

where $\rho(x, y)$ is the distance of the ith and j^{th} edge cell from all other cells in the grid. η is a scaling factor dependent linearly on L_{Road}. The calculated values are ceiled to an appropriate value. The total edge potential is a summation of potential values of all edge-containing cells.

3.1.2 Road potential

Based on a sociological study conducted in France [16], it can be inferred that the propensity of illegally crossing a road is linearly dependent on its width. Thus, a narrow road entices a pedestrian to cross illegally while a wider one does not.

For cell (C^{ij}) with an attribute *road*, the calculated potential value is:

$$\mathbf{U}^{ij}_{Road} = \beta_{Road}\exp\left(-\left[\left(\frac{x_{ij}-x_{road}}{\sigma_x}\right)^2 + \left(\frac{y_{ij}-y_{road}}{\sigma_y}\right)^2\right]\right) \tag{3}$$

β_{Road}, σ_x and σ_y are dependent on the width of the road as explained earlier.

The total road potential is the summation of potential values of all road-containing cells.

3.1.3 Obstacle Potential

Obstacles in the scene can be distinguished as static and dynamic obstacles. For either classification, the response of the self-driven particle under the effect of the obstacle remains the same. The self driven particle cannot cross through the obstacle and the approach to the obstacle is slow. A Yukawa potential [17] is considered a fit for the expected behaviour.

A static obstacle takes the shape that it is perceived to be. A dynamic obstacle(for example, other cars in the scene), is described as a rectangular shape with a triangular shape extending forward in the direction of motion. Thus, the potential is described by:

$$\mathbf{U}^n_O = \Lambda\frac{\exp\left(-\alpha\mathbf{K}\right)}{\mathbf{K}} \tag{4}$$

Where Λ and α decide the behaviour of U^n_O. Larger the values, sharper the drop off of the potential near the obstacle.

$\mathbf{K}$ is the distance of the obstacle from every point on the workspace, i.e.,

$$\mathbf{K} = \|C^{ij} - C^{Obs}\| \tag{5}$$

The total effect of all the obstacles in the workspace is given as

$$\mathbf{U_{Obs}} = \sum_{n=0}^{N}\mathbf{U}^n_O \tag{6}$$

Where N is the total number of obstacles observed. The extremely large values that are generated are truncated to a maximum viable value.

3.1.4 POI Potential

A Point of Interest (an inexhaustive list of what may be considered as a POI can be found in [5]) generates an attractive pull in the scene. With sufficient motivation, the self-driven particle can escape the influence of a POI. A POI is also a terminal point in the scene - the implication being that all exits in the scene are POIs. The potential of a POI situated at a cell defined by (x_{poi}, y_{poi}) is a Gaussian function centered at (x_{poi}, y_{poi}). $\beta_{poi}, \sigma_x, \sigma_y$ depend on the global importance of the specific Point of Interest.

3.1.5 Cross-walk Potential

The cross-walk connects the two side-walks of the street and acts as a resistance-less conduit for the self driven particle. Thus, the potential of the cross-walk is the smallest value in the area.

3.2 Behavioural Classification

3.2.1 Trajectory Regions

Pedestrian route choice behaviour, in inner city limits, can be described in terms of optimisations. A pedestrian either tries to perform a distance optimisation to a destination at one extreme or optimises for safety at the other. Thus, all possible pedestrian paths can captured between these two behaviours. From each destination in the scene to all the others, an A* search is performed with the heuristic:

$$h(s) = \alpha C(s, s+1) + (1-\alpha)\rho(g, s+1) \tag{7}$$

Where $C(s, s+1)$ is the cost of moving from the current state s to the next state $(s+1)$. $\rho(g, s+1)$ is the normalised distance between the goal (destinations) and the next state. α is the parameter contributing to the integration of safe trajectory optimisation and shortest distance optimisation. Varying this parameter allows for simulating the different trajectories pedestrians might take, like partially optimising for distance and partly for safety.

3.2.2 Probability map of pedestrian trajectories

Trajectories generated for each entrance and exit are collated to create regions of probable pedestrian trajectories. These are then analysed to find regions of overlap. The larger the value of the region of overlap, higher the probability that a pedestrian might be present there.

4 IMPLEMENTATION AND RESULTS

In the current work, the red block in Fig. 1 has been implemented on the dataset provided in [18]. This is a dataset generated in Martigny, Switzerland from a static camera overlooking a city square. The camera captures a scene containing 2 POIs and a crosswalk. It also captures pedestrian and traffic circulation. POI positions have been extracted from OpenStreetMaps. The scene at a specific instant is shown in Fig. 2

This scene is segmented into a grid containing features of the component potential fields - road, POI, crosswalk and sidewalks. Road parameters are taken from Swiss national standards and a potential field resulting from the constituent components are created as explained in 3. This potential field "Map" is then used to determine areas of probable trajectories and pedestrian probability.

4.1 Validation

To show that this approach follows *Natural Vision* [4] is sufficient validation to use it to efficiently model the environment. Naive and sufficient validations are described for the approach, based on chosen destinations.

Fig. 4 forms the crux of our work. This scene, taken from the acquired dataset, is the representation of the observation at a specific instant (leftmost image in Fig. 4). The width of the road and lanes were extracted manually based on Swiss national standards. With this data, the scene is reconstructed by placing its constituent elements at relative positions on a grid. The obstacle has been identified and tracked using the YOLO2 framework. Attributes of each cell on the grid is manually defined as crosswalk, side-walk, road, edge or obstacle, as well as for the POI. With these data populated on the grid, a resultant potential field "Map" has been generated as defined in section 3.

4.1.1 Primary Validation

A first naive validation of this method is to show that observed pedestrian presence matches those areas predicted by the framework. In this scene (Fig. 2), there are 4 potential destinations, marked (1), (2), (3) and (4). Destination (2) is the entrance of the visible POI while the others are the ends of the scene. Choosing one of these as a starting point, we compute trajectories of the safest and the shortest paths and everything in between to all the other exits. By repeating this at each exit, a map of areas of most probable pedestrian positions emerges. For instance on Fig. 3, two destinations – (1) and (4) – were chosen and the different trajectories determined. Superimposing areas leads to the pedestrian trajectory probability map at right on the same figure.

Pedestrians were detected using the YOLOv2 framework [19]. When the YOLOv2 framework failed, there was a manual annotation of pedestrians for detections. Trajectories were tracked based on these

Figure 2: Scene at a specific instant from the dataset. No dynamic obstacles on the road are observed at this instant.

detections. By accumulating these trajectories, a probability map based on observations was created (Section 3.2.2). Fig. 3 compares the predicted probability areas from our model and the observation.

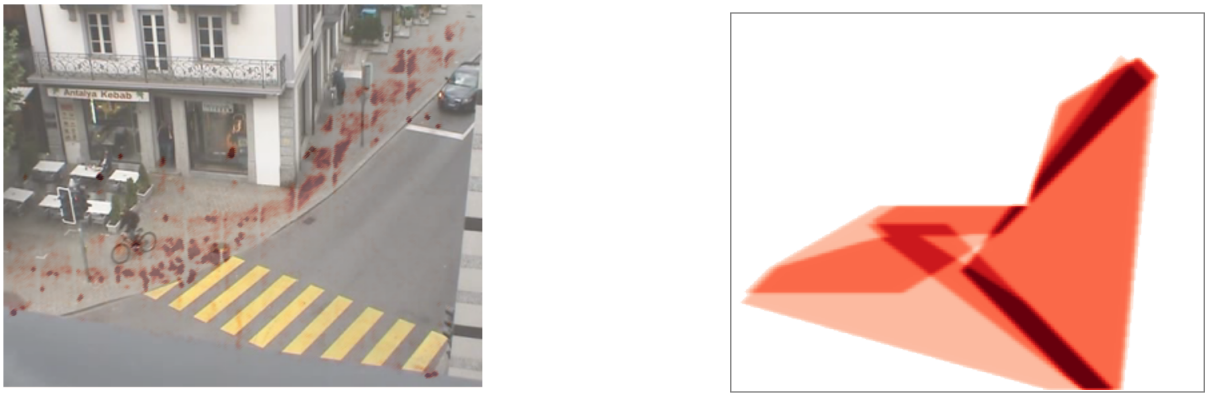

Figure 3: Comparison between observed probability map and predicted pedestrian trajectory probability map. The left image shows the ground truth. Right image is the predicted probability map of the scene. Darker the colour, higher the probability of pedestrians being present in that area.

In the observed probability map, it can be seen that there is a high incidence of pedestrians on the side-walk. The probability of a legal crossing is much higher compared to an illegal one, as observed. The predicted pedestrian trajectory probability map shows that there is a high probability that this scene has more chances of pedestrians crossing legally compared to an illegal crossing. It also predicts a high number of pedestrian trajectories into destination (3). Comparing to the ground truth, even though there are a few stray illegal crossings, they are overwhelmingly outnumbered by the number of legal ones. It is also observed that pedestrians continue to walk towards destination (3) compared to people crossing across into destination (2) in the scene.

4.1.2 Secondary Validation

A secondary validation of the framework is to prove that it predicts behaviour based on Natural Vision. Pedestrians stay within certain bounds while going towards a destination. Knowing this, areas of illegal crossings can be predicted. From the dataset, 38 trajectories were chosen and classified into eight different behaviours. Each case is classified based on the entry and exit destinations of the scene and the legality

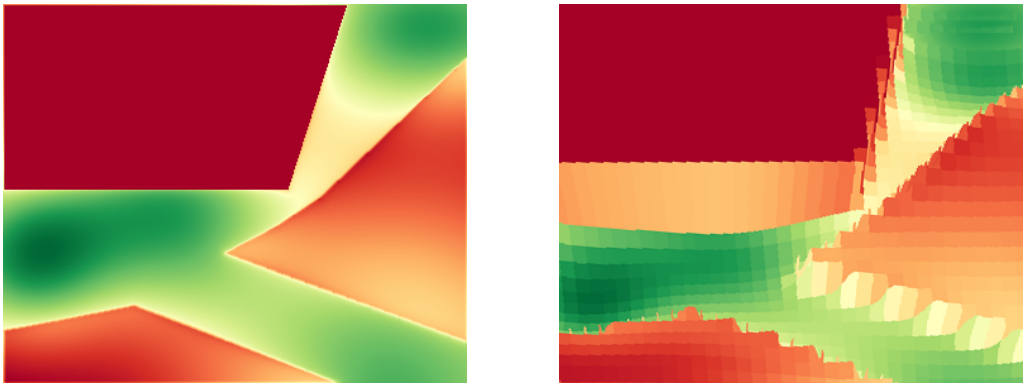

Figure 4: Resultant potential field of the observed scene. The scene contains a visible POI, a total of four destinations and a cross-walk. The right image is the 3D representation of the same scene. At this instant, there are no dynamic obstacles observed, the presence of which, changes the resultant field.

Table 1: Different behavioural cases observed in the dataset. Destinations can be seen on Fig 2.

Case	**Destinations**		**Crossing**
	From	**To**	**Legality**
I	1	3	Legal
II	1	2	Legal
III	4	2	Legal
IV	4	1	Legal
V	1	4	Illegal
VI	1	4	Legal
VII	4	1	Illegal
VIII	4	3	Legal/illegal

of crossing between them. The behavioural classes for pedestrians walking in the scene (Fig. 2) can be found in Table 1.

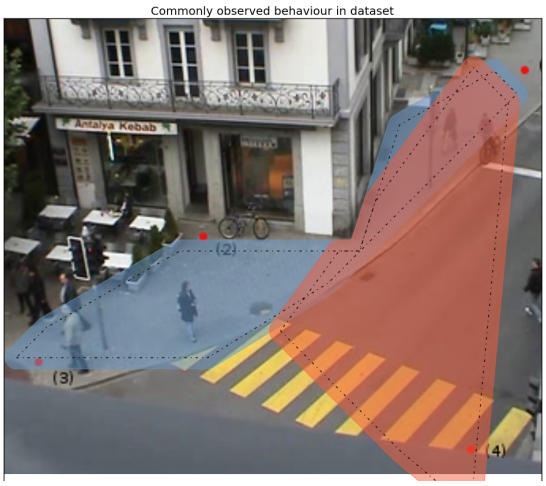

Figure 5: Validation of the two most common behaviours in the dataset: case I and case VI (see table 1). The dashed line corresponds to the A* predicted regions for each case. The coloured boundaries represent the extended zone.

As seen in section 3.2.1, the A* algorithms returns a sharp trajectory, going straight for the goal based on the heuristic provided. It does not meander when there are no explicit potential modifications. The zones predicted by just using the A* algorithm can be seen as the dashed line in Fig. 5. Human motion very rarely follows a perfect straight line. Thus, to account for these random motions, the zones were dilated by 40 cms on all sides. This leads to a much better prediction score as can be seen in Table 2. This is sufficient to show that the principle of *Natural Vision* is valid and can be used to determine pedestrian behaviour.

Table 2: Quantitative analysis of trajectories within predicted regions

Case	Nb Trajectories	A* Predicted zone		Extended zone (40cms)	
		Inside zone (%)	Outside zone (%)	Inside zone (%)	Outside zone (%)
I	10	84.11	15.88	96.88	3.11
II	2	30.07	69.92	88.17	11.82
III	1	29.10	70.89	51.49	48.50
IV	9	68.48	31.51	77.43	22.56
V	7	72.38	27.61	83.26	16.73
VI	9	39.63	60.36	50.26	49.73
VII	12	76.79	23.20	83.28	16.71
VIII	10	64.67	35.32	69.98	30.01

4.2 Discussion

A very high percentage of the most common behaviours observed in the dataset, I and VII, are seen to be within the predicted zone, regardless of extending the zone. For some behavioural cases, like cases II and III, not many pedestrian trajectories could be found in the chosen dataset. These zones are bounded by the shortest route to the goal and the safest. Our results are a validation of this. These scores can be ameliorated by substituting better parameters to build the potential field model. For example, the *attractiveness* parameter of the POIs in the scene were assumed to be equal for all POIs in the scene. These parameters could change based on the POIs global importance, the time of day, etc., all of which could be encoded on a map for a given city. By accurately estimating the values, pedestrian behaviour in urban centers could be much better predicted. Thus, a conclusion can be reached that pedestrian behaviour is not random in nature. Their movement can be accurately predicted by utilising well established sociological ideas of *attractors* and *Natural Vision* as our work demonstrates. This helps in providing prior knowledge of the observed scene which could then be used for predicting pedestrian intentions. An effect of the primary validation is that the resultant pedestrian trajectory probability map could provide prior knowledge of hazardous zones in an urban street that can be used for safer navigation.

5 CONCLUSION

In this work, we have established a new framework for increasing the Situational Awareness of an autonomous car on urban roads. This is done by adapting the sociological principle of Natural Vision as a function of a potential field composed of different elements of the urban environment. This allows the car to understand pedestrian behaviour in previously unobserved areas. We have also demonstrated that pedestrian behaviour in urban areas is not random but is a function of the built environment they are in. The main contributions of this paper are – a) the usage of sociological principles and the integration of POIs into understanding pedestrian behaviour, b) quick computation of probable pedestrian movement zones even when there are no pedestrian observations in the scene.

The current work utilises a dataset that captures pedestrian behaviour with a mounted, stationary camera overlooking a single view. Future work will deal with the application of this framework from the ego-perspective of an autonomous car. Algorithms like the CMC-DOT [20] can be used for estimating the occupancy grid and positions of dynamic obstacles in real time. This can then be used as an input for our framework. Another work that needs to be done is to have online extraction of POIs and track individual pedestrians based on the priors determined by our framework.

ACKNOWLEDGMENT

These researches have been conducted within the VALET project, funded by the French Ministry of Education and Research and the French National Research Agency (ANR-15-CE22-0013-02).

References

[1] L. Gugerty, "Evidence from a partial report task for forgetting in dynamic spatial memory," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 40, no. 3, pp. 498–508, 1998.

[2] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 37, no. 1, pp. 32–64, 1995.

[3] B. Hillier, A. Penn, J. Hanson, T. Grajewski, and J. Xu, "Natural movement: or, configuration and attraction in urban pedestrian movement," *Environment and Planning B: planning and design*, vol. 20, no. 1, pp. 29–66, 1993.

[4] J. J. Gibson, "The ecological approach to visual perception." 1979.

[5] D. Helbing and P. Molnar, "Social force model for pedestrian dynamics," *Physical review E*, vol. 51, no. 5, p. 4282, 1995.

[6] E. Papadimitriou, G. Yannis, and J. Golias, "A critical assessment of pedestrian behaviour models," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 12, no. 3, pp. 242–255, 2009. [Online]. Available: http://dx.doi.org/10.1016/j.trf.2008.12.004

[7] K. Yamaguchi, A. C. Berg, L. E. Ortiz, and T. L. Berg, "Who are you with and where are you going?" in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 1345–1352.

[8] S. Pellegrini, K. Schindler, and L. van Gool, "You'll Never Walk Alone: Modeling Social Behavior for Mluti-target Tracking," *Proceedings of the IEEE 12th International Conference on Computer Vision (ICCV)*, no. Iccv, pp. 261–268, 2009.

[9] T. Bandyopadhyay, C. Z. Jie, D. Hsu, H. Marcelo, A. Jr, D. Rus, and E. Frazzoli, "Intention-Aware Pedestrian Avoidance," *The 13th International Symposium on Experimental Robotics*, pp. 963–977, 2013.

[10] B. D. Ziebart, N. Ratliff, G. Gallagher, C. Mertz, K. Peterson, J. A. Bagnell, M. Hebert, A. K. Dey, and S. Srinivasa, "Planning-based prediction for pedestrians," in *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*. IEEE, 2009, pp. 3931–3936.

[11] K. Kitani, B. Ziebart, J. Bagnell, and M. Hebert, "Activity forecasting," *Computer Vision–ECCV 2012*, pp. 201–214, 2012.

[12] D. Vasquez, "Novel planning-based algorithms for human motion prediction," in *Robotics and Automation (ICRA), 2016 IEEE International Conference on*. IEEE, 2016, pp. 3317–3322.

[13] S. Bonnin, T. H. Weisswange, F. Kummert, and J. Schmuedderich, "Pedestrian crossing prediction using multiple context-based models," *2014 17th IEEE International Conference on Intelligent Transportation Systems, ITSC 2014*, pp. 378–385, 2014.

[14] O. Khatib, "Real-time obstacle avoidance for manipulators and mobile robots," *The International Journal of Robotics Research*, vol. 5, no. 1, pp. 90–98, 1986.

[15] M. T. Wolf and J. W. Burdick, "Artificial potential functions for highway driving with collision avoidance," in *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*. IEEE, 2008, pp. 3731–3736.

[16] M. C. Montel, T. Brenac, M.-A. Granie, M. Millot, and C. Coquelet, "Urban environments, pedestrian-friendliness and crossing decisions," in *Transportation Research Board 92nd Annual Meeting*, 2013, p. 13p.

[17] R. Volpe and P. Khosla, "A theoretical and experimental investigation of impact control for manipulators," *The International Journal of Robotics Research*, vol. 12, no. 4, pp. 351–365, 1993.

[18] J. Varadarajan and J.-M. Odobez, "Topic models for scene analysis and abnormality detection," in *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*. IEEE, 2009, pp. 1338–1345.

[19] J. Redmon and A. Farhadi, "Yolo9000: Better, faster, stronger," *arXiv preprint arXiv:1612.08242*, 2016.

[20] L. Rummelhard, A. Nègre, and C. Laugier, "Conditional monte carlo dense occupancy tracker," in *Intelligent Transportation Systems (ITSC), 2015 IEEE 18th International Conference on*. IEEE, 2015, pp. 2485–2490.