
From Pedestrian Detection to Crosswalk Estimation: An EM Algorithm, Analysis, and Evaluations on Diverse Datasets

Ross Greer¹ Mohan Trivedi¹

Abstract

In this study, we contribute an EM algorithm for estimation of corner points and linear crossing segments for both marked and unmarked pedestrian crosswalks using the detections of pedestrians from processed LiDAR point clouds or camera images. We demonstrate the algorithmic performance by analyzing three real-world datasets containing multiple periods of data collection for four-corner and two-corner intersections with marked and unmarked crosswalks. Additionally, we include a Python video tool to visualize the crossing parameter estimation, pedestrian trajectories, and phase intervals in our public source code.

1. Introduction

The detection of crosswalks serves as an important module in pedestrian detection and tracking systems for both intelligent vehicles and smart infrastructure, especially when it comes to safety, as accidents involving pedestrians is one of the leading causes of death and injury around the world (Gandhi & Trivedi, 2008a).

The expected behavior of a vehicle (autonomous or not) in such situations is made explicitly clear by the California DMV Driver Handbook: “If you approach a pedestrian crossing at a corner or other crosswalk, even if the crosswalk is in the middle of the block, at a corner with or without traffic signal lights, whether or not the crosswalk is marked by painted lines, you are required to exercise caution and reduce your speed, or stop if necessary, to ensure the safety of the pedestrian.”

Autonomous vehicles adhere to this policy and reduce speed when a crosswalk is detected; such control methods are

described in (Zhu et al., 2021), also accounting for the uncertainty in occluded scenes. Driver assistance systems have also been designed to encourage similar slowing; for example, the detection of a crosswalk activates a pedestrian collision warning system in (Suzuki et al., 2010). Modules and data supporting these practices are extensive; as examples, the crowd-detection system sensitivity in (Reisman et al., 2004) is tuned based on the presence of a crossing, and driver slowing patterns at crosswalks and associated pedestrian behaviors have been thoroughly studied in (Zang et al., 2021) and (Robert et al., 2021), providing a useful model for autonomous vehicles to adopt. Further, autonomous vehicles exercise caution to ensure safety of pedestrians through the estimation of pedestrian paths, which can be enhanced by the presence of crosswalks which can be used to estimate vehicle risk of encountering a pedestrian (Tian et al., 2014) and suggest a higher probability of pedestrian crossing behavior as well as a prescribed motion (Gandhi & Trivedi, 2008b). As some motivating examples, (Chen & Zhang, 2022) shows that knowledge crosswalk boundaries improve this path prediction using a modified social force model, and (Quintero Mnguez et al., 2019) explains that spatial layout of the environment, including crosswalks, provides valuable information to Automatic Emergency Braking Systems. (Deo & Trivedi, 2017) and (Kitani et al., 2012) use scene information such as crosswalks, origins, and destinations to learn a prior distribution on future pedestrian trajectories, and this information appears repeatedly in the research in the literature review on pedestrian behavior prediction presented in (Ridel et al., 2018). The principle is further evident in (Zhu et al., 2022), which shows that contextual information is particularly useful in trajectory prediction when using the technique of ‘unlikelihood training’, where trajectory outputs which violate contextual expectations are assigned lower probability. (Fridovich-Keil et al., 2020) shows that knowledge of goals not only informs trajectory predictions, but more importantly can properly adjust an autonomous system’s expectations in its prediction confidence when a pedestrian behaves in a way which deviates from direct paths to its goal; modeling this uncertainty can help autonomous agents perform more safely in human interactions. Probabilistic compact representations of the on-road environment which encapsulate spatial, dynamic, and legal

¹Laboratory for Intelligent Safe Automobiles, University of California, San Diego, USA. Correspondence to: Ross Greer <regreer@ucsd.edu>.

information have been successfully applied in previous research, for example, in facilitating safe lane changes and merges (Sivaraman & Trivedi, 2014). In general, estimation of crosswalk parameters from outside-facing cameras belongs to the problem posed in (Satzoda & Trivedi, 2014); that is, extracting semantics of interest from the surrounding scene for the purpose of analyzing related behaviors of pedestrians, drivers, and autonomous vehicles.

Awareness of crossings is of particular importance in high-speed environments, where street cross markings can lure pedestrians into a false sense of security when drivers are unaware of the crossing (Gandhi & Trivedi, 2007).

In addition to benefits of pedestrian and driver safety, even the vehicle's own path planning objectives and related tasks can benefit from accurate crosswalk detections; as an example, vehicle localization can be longitudinally corrected based on crosswalk estimates (Jo et al., 2013).

Therefore, we propose an algorithm which, given a set of 3D pedestrian detections, estimates lines which model the de facto pedestrian crosswalks at an intersection. Figures 1 and 2 illustrate the purpose of the algorithm within the context of connected safety systems, where an accurate representation of scene information can improve effectiveness of vehicle safety measures taken at crossings by inferring marked and unmarked pedestrian paths to provide stronger priors for trajectory prediction.

2. Related Research

Past approaches to detecting crosswalks have been primarily based on camera images of the static scene. Such methods include figure-ground segmentation (Coughlan & Shen, 2006), edge direction and cross-ratio analysis (Suzuki et al., 2010), Canny edge detection and Hough line parameter estimation (Tian et al., 2021), MSER and ERANSAC (Zhai et al., 2015), laser scanning (Hernandez et al., 2015), monocular image deep learning to classify presence of crosswalk in overall scene (Tmen & Ergen, 2020), PCA-based road segmentation and crosswalk detection from a LiDAR point cloud (Riveiro et al., 2015), and image classification using deep learning over massive instances from satellite imagery (Berriel et al., 2017).

Camera-based approaches are the most popular for such detections, and for good reason: (Song et al., 2022) shows that in urban point cloud maps, crosswalks and other road surface markings (RSMs) are the most challenging class in point cloud semantic segmentation. But even in camera-based approaches, cracks and other damage to painted lines can provide significant challenges, with remedies proposed in (Ito et al., 2021).

Rather than modeling the scene based using computer vision

applied to static scene elements, we are the first to use the motion patterns of detected pedestrians to estimate the crosswalks. Our approach provides a few unique benefits:

- The algorithm can directly recognize crosswalks even under the additional DMV stipulation “whether or not the crosswalk is marked by painted lines”. Pedestrians following such trajectories are not necessarily illegally jaywalking; they are simply using an unmarked crosswalk.
- The algorithm parameterizes crosswalks based on actual crossing paths and rest points, which may differ from those indicated by prescribed road markings.
- No prior knowledge of the intersection geometry is required, only pedestrian detections.
- The algorithm is robust to non-square, non-rectangular, and even non-parallelogram intersection-crossing geometries.
- The algorithm is robust to natural degradation of road markings (chipped paint, graffiti) or temporary detours to paths (e.g. construction).

3. EM Algorithm for Crosswalk Estimation

We make the following assumptions about pedestrian behavior:

1. The crosswalk being utilized by a pedestrian p_i is a latent variable z_{p_i} on which their motion depends.
2. Pedestrians spend, on aggregate, more time at a crossing corner than on any point of the street crossing. This enables the use of k-means clusters for initialization of crossing corners.
3. Crosswalks are linear in nature. This enables the estimation of crossings as lines of best-fit.
4. Most pedestrians cross within the designated or de facto crossing bounds with an average trajectory along the middle of the crosswalk. This enables the filtering of outlier crossing points.

The Expectation-Maximization (EM) algorithm (Dempster et al., 1977) is an iterative algorithm used to find maximum-likelihood estimates of parameters of a statistical model in the presence of latent (hidden) variables. We frame our statistical model to be the distribution of pedestrian detections, subject to the parameters which define the location of street corners. These corners are parameters which dually define the crosswalks as the connecting line segments. Samples from the distribution of the pedestrian detections depends on

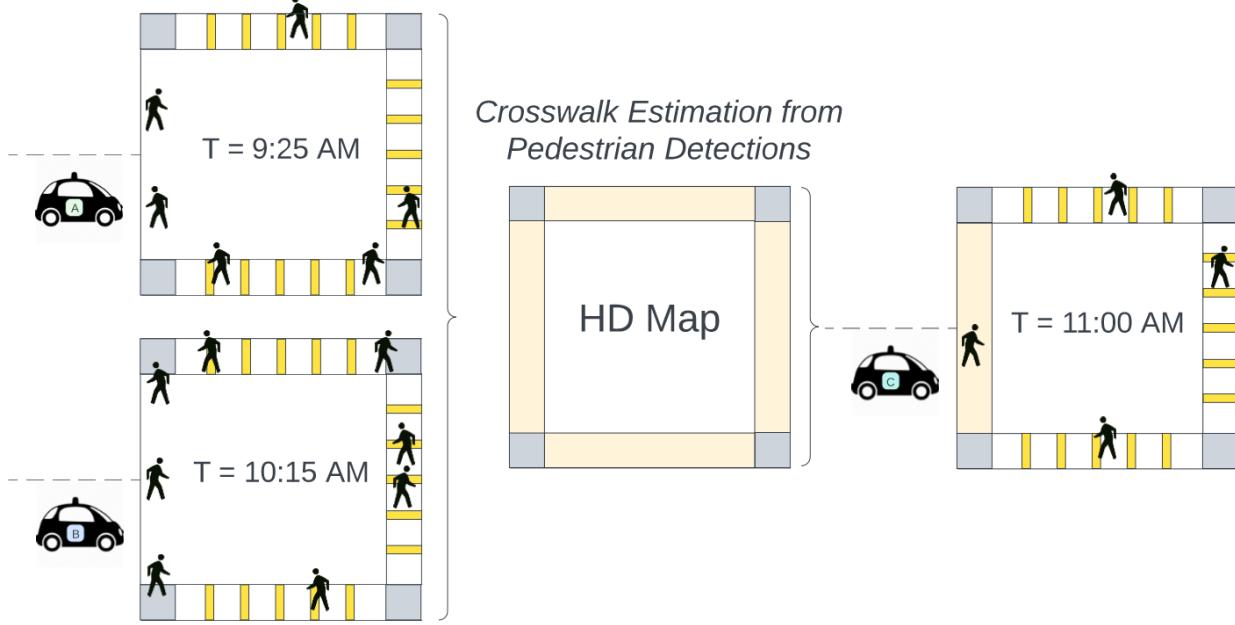


Figure 1. In this scene, two different intelligent vehicles approach the same intersection at different times, detecting pedestrians on both marked and unmarked crosswalks. The detections are provided to the crosswalk estimation algorithm proposed in this paper. The output crosswalk parameters are used to update an HD map. Intelligent vehicles approaching the same intersection now have information about the presence of the unmarked crosswalk, which can be employed in Automatic Emergency Braking Systems or similar systems for pedestrian safety and comfort, as well as systems designed to predict and classify pedestrian trajectories.

a latent variable: which crosswalk is the detected pedestrian occupying? Given input of 3D pedestrian detections within a coordinate system, the following EM algorithm estimates the locations of street corners which act as starting or ending points for crosswalks (marked or unmarked). An overview of the algorithm is shown in Figure 3.

1. Initialization

- Begin with (x, y) ground plane positions for all pedestrian detections made using LiDAR points.
- Perform k-means clustering (Likas et al., 2003) where k is the number of corners expected at the intersection. In the case of our dataset, $k = 4$. These cluster centers will serve as preliminary corners.
- Compute the angle between each pair of corners.
- Using the corners central to the greatest four angles, form four pairs (c_n, c_m) of corners, which parameterize a given crossing segment. The greatest angles are used to ensure that these segments comprise the border of the intersection rather than diagonal segments which interfere with traffic. In the case of environments where diagonal crossing is expected, additional segments can be allowed.

2. Expectation

- Assign to each pedestrian detection p_i an expected value of the latent variable z_{p_i} referring to the associated crossing line l_j , parameterized as (a_j, b_j, c_j) . In this case, we compute this nearest crosswalk value using

$$z_{p_i} = \arg \min_j \frac{|a_j x_{p_i} + b_j y_{p_i} + c_j|}{\sqrt{a_j^2 + b_j^2}} \quad (1)$$

- Filter by ignoring outlier points which are more than M meters from all crossing lines during this EM iteration. M is a hyperparameter; we use a value of 3.5 meters.

3. Maximization

- Fit a line of least-squares $l_{z_{p_i}}$ to each set of detections associated with a unique z_{p_i} , representing the crossing. It is important to note that initial estimates are easily trapped in local minima (for example, if k-means initializes near a non-corner position that a pedestrian chose to stand in wait). To allow the estimate to escape this local well, we use two remedies:

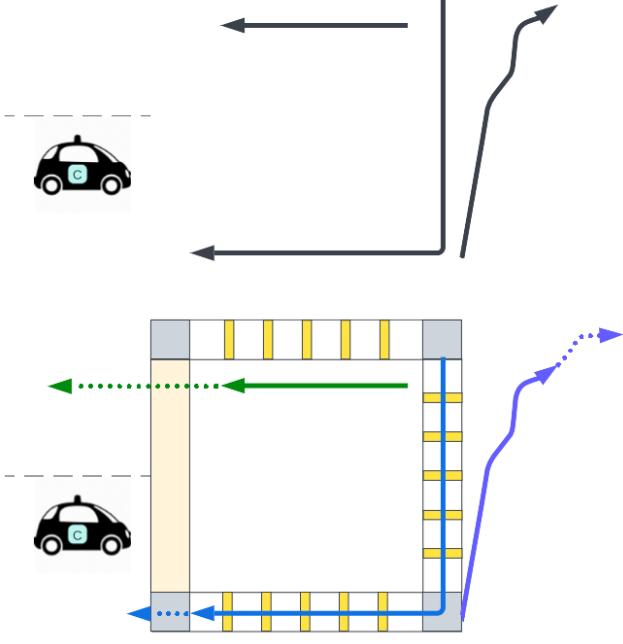


Figure 2. Understanding pedestrian scene infrastructure parameters provides stronger information for the intelligent vehicle to classify the pedestrians, which in turn informs possible future trajectories. In this illustration, the pedestrian in green is riding through a bike lane, in blue reaching a corner destination, and in purple jaywalking to a sidewalk.

- Treat the detections as an occupancy map rather than a heatmap (that is, each location within some resolution can be either occupied or non-occupied, rather than allowing multiple detections to be representative of the same area). This is useful in reducing the effect of a person standing in a non-corner area for a long period of time, while still capitalizing on the density of points near corners from regular pedestrian traffic. We use a resolution of 0.1 meters.
- Begin the estimate with a line estimate method which is less sensitive to outliers (in our case, we use the Theil-Sen Estimator (Pedregosa et al., 2011)).

Both of these modifications can be dropped once the intersection corners make an initial convergence within some tolerance t_1 .

- Define two parallel boundary lines surrounding each crossing, with distance offset

$$d = \max(d_u - \frac{n}{r_m}, d_l), \quad (2)$$

where n is the iteration number, d_u and d_l are

the upper and lower limits to the margin of allowable pedestrian points to be considered ‘on’ the nearest crosswalk, and r_m is the rate at which these margins narrow, in meters per iteration. To find these lines, first compute the angle of inclination $\alpha = \arctan m$, where m is the slope. The complement of α subtends a right triangle whose opposite side is d and hypotenuse is b' , a term which can be added to the initial y-intercept of the line to form a parallel line offset at distance d . Note that this step should be skipped until t_1 is satisfied; else the restrictive margins may ‘race’ against the initial convergence of corners to a space away from outliers.

- Filter points outside the boundaries from the next iteration. As n increases, the boundaries tighten to include only the most representative points of the crossing. This also reduces the effects of trajectories deviating from the crossing lines after the pedestrian’s segment between the corners is complete.
- Repeat E and M steps until convergence. To determine stopping criteria, take the intersection of each pair of lines as the corresponding crossing corner, and terminate when the corner positions do not change by more than tolerance t_2 in sum. Both t_1 and t_2 are hyperparameters; we use values of 0.3 and 0.05 meters.

4. Analysis

4.1. Datasets

Our experiments utilize three datasets. First, we utilize the data presented in the 2022 Transportation Review Board Committee on Artificial Intelligence and Advanced Computing Applications Transportation Forecasting Competition (Ishak et al., 2021). Three time periods of data were collected from three Ouster OS1-128 LiDAR sensors at the MLK and Georgia Avenue intersection in Chattanooga, Tennessee (see Figure 4). The LiDAR point cloud data was processed by Seoul Robotics software to provide object labels (pedestrian, vehicle, and bicyclist), and sub-classification features of object size, position, and velocity. Each pedestrian data entry also includes a confidence and tracking status. Second, we use pedestrian detection data collected from the Stanford Drone Dataset (Robicquet et al., 2016) at a bicycle traffic circle 5, where pedestrians exhibit similar behavior to street crossing. Note that this dataset provides detections on a pixel scale rather than meters. Third, we use 3D pedestrian detection data collected from a two-corner crosswalk on the campus of University of California, San Diego, collected from a camera mounted to a stationary parked vehicle (see Figure 6) (Deo & Trivedi, 2017).

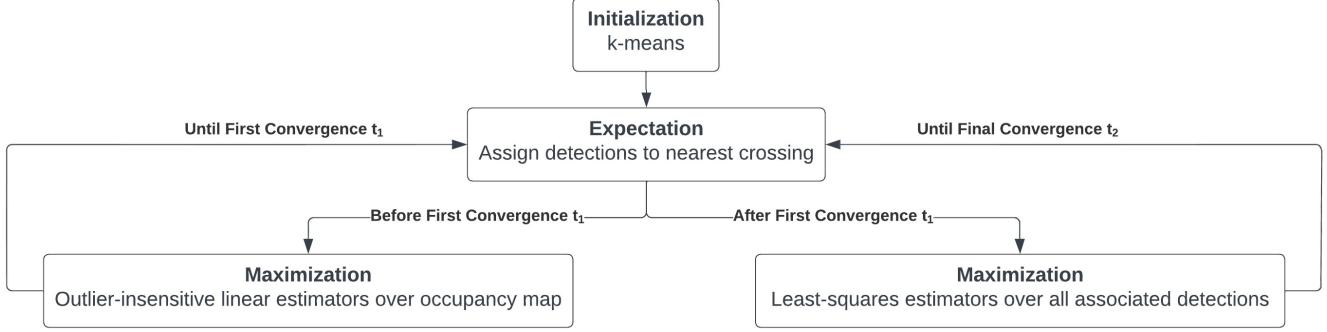


Figure 3. Proposed Algorithm for Crosswalk Estimation



Figure 4. Ground and bird's-eye-view of the intersection of MLK and Georgia Avenue in Chattanooga, Tennessee, where infrastructure-mounted LiDAR pedestrian detections were captured.

The assumption that pedestrian detections tend to aggregate at crossing corners is valid for the analyzed datasets. This property is illustrated in the data collected at the intersection in Chattanooga, shown in Figure 7.

4.2. Results

4.2.1. FOUR-CORNER CROSSING: CHATTANOOGA

Figures 8, 9, and 10 illustrate the application of the above EM algorithm to estimate the location of crossing corners for each of the three data collection periods. Observing the coverage of the pedestrian detections, it is evident that the final EM algorithm estimates (green) are much closer to the true crosswalk corners than the initial k-means estimates (red). Notably, the EM algorithm is able to move the corner estimates to areas less densely populated with detection points to improve accurate positioning, and is robust to the many noisy data points which exist in non-crossing regions.

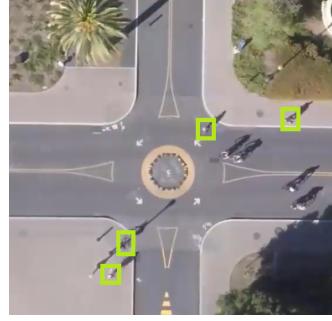


Figure 5. Bird's-eye-view of the traffic circle in Palo Alto, California, from the Stanford Drone Dataset. It is notable that this intersection contains no marked crosswalks.

Despite the three samples containing completely different detection datapoints and k-means initializations, the corners converge to similar locations as described in Table 1. The EM algorithm shows consistent performance in providing an estimate near the true corner locations, especially on the first and third batches where the initialization is very distant from the goal. In the case of Batch 2, the initialization was already very close to the corners, and it is worth reminding that this algorithm does not necessarily converge to the marked street corners (as approximated by human annotation) as the estimates reflect the actual positions which tend to be occupied by pedestrians; in this batch, we can see that pedestrians tended to aggregate close to the inner portion of the corner.

Table 1. Average distance (m) of batch estimate from labeled ground truth corners using k-means initialization alone vs. to complete EM algorithm. Ground-truth corners were labeled by a human annotator.

| | Batch Size | k-means (m) | EM (m) | % Improve |
|---------|------------|-------------|--------|-----------|
| Batch 1 | 13,054 | 8.884 | 1.192 | 86.6 |
| Batch 2 | 61,918 | 1.385 | 1.715 | -23.8 |
| Batch 3 | 19,033 | 2.698 | 1.158 | 57.1 |



Figure 6. Ground and bird's-eye-view of the crosswalk in La Jolla, California, where vehicle-mounted camera pedestrian detections were captured.

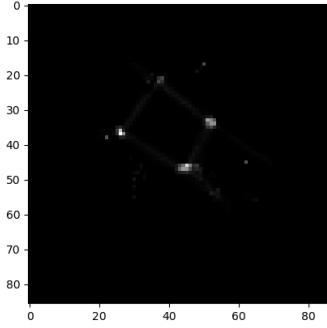


Figure 7. A heatmap visualizing pedestrian detection instances on a grid over the Chattanooga intersection. Brighter pixels indicate more pedestrian data instances belonging to that region. From the figure, we can see that the most points occur at the four corners, where a pedestrian would be waiting to cross, with faint lines over the pedestrian crosswalks.

The first batch (Figure 8) shows the most dramatic shift from initial to final locations. This data period was heavily skewed by pedestrian behavior which occurred down the road from the intersection, but the algorithm was able to overcome the initial local minimum through the important step of occupancy-map formulation and outlier-insensitive linear estimation.

The corresponding crosswalks are illustrated in Figure 11, including a margin inside and outside each line to account for pedestrian deviations from the center crossing line. Depending on intended application, this margin can be provided as a hyperparameter to match safety or legal expectations, or learned from the data to match practical usage. For example, in our illustration, we use values for inner and outer margins of $I = 2$ and $O = 2.75$ respectively. The greater tolerance for O is employed since pedestrians can safely wait further from the corner than the sidewalk, but

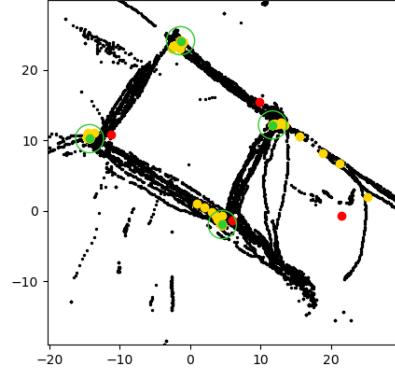


Figure 8. Estimation of street corners associated with crosswalks for the first batch of data collected in Chattanooga. Pedestrian positions are represented in black, initial cluster centers estimating corners in red, progression of the corners between iterations in yellow, and final corner positions in green.

cannot safely wait within the street.

Using the estimated corners and crosswalks, we show that it is possible to classify pedestrian trajectories using a rule-based technique (i.e. pedestrians passing travelling from corner-to-corner are legally crossing, pedestrians who move through an intersection via points adjacent to but not upon corners are likely bikers, pedestrians who cross from a sidewalk to another sidewalk but avoid the corners are jaywalking, etc.). These classifications, shown in Figure 12, have value in predicting future trajectories of the pedestrians for safe vehicle operation.

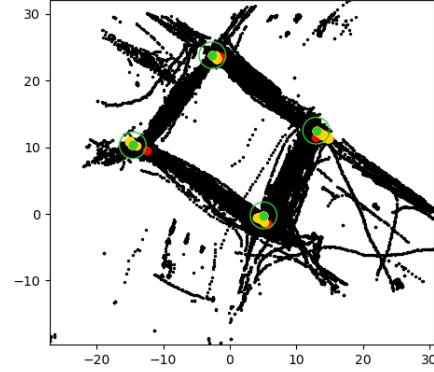


Figure 9. Estimation of street corners associated with crosswalks for the second batch of data collected in Chattanooga, following the same color scheme as described in Figure 8.

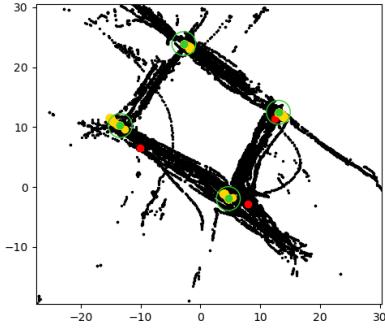


Figure 10. Estimation of street corners associated with crosswalks for the second batch of data collected in Chattanooga, following the same color scheme as described in Figure 8.

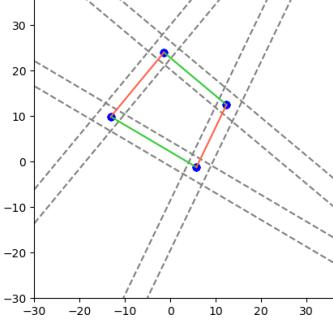


Figure 11. Model representation of the intersection. Corner points are depicted in blue, representing the average of the corners found across the three data batches. In this figure, to best illustrate the scene, parallel crossings are shown as solid green lines due to an active crossing phase, while the other set of parallel crossings are shown in red to indicate non-active crossing phase. Dashed lines indicate our defined inner and outer boundaries for each crossing.

4.2.2. FOUR-CORNER CROSSING: STANFORD TRAFFIC CIRCLE

We repeat the proposed algorithm on a bike traffic circle from the Stanford Drone dataset. Most notably, this consists entirely of unmarked crosswalks, yet the pedestrian trajectories still produce a clear pattern which is otherwise invisible in road infrastructure. The k-means initialization places the estimates in locations which are near the corners but not precise; after 27 iterations, the corners converge to positions which represent the most frequent (unmarked) crosswalks used by the pedestrians.

4.2.3. TWO-CORNER CROSSING: LA JOLLA

Using the modification of the algorithm for a single crosswalk (no intersecting corners), we demonstrate the algo-

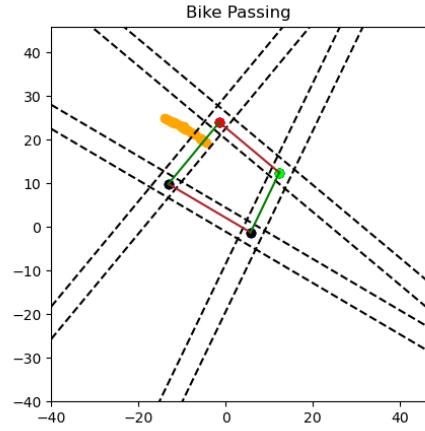
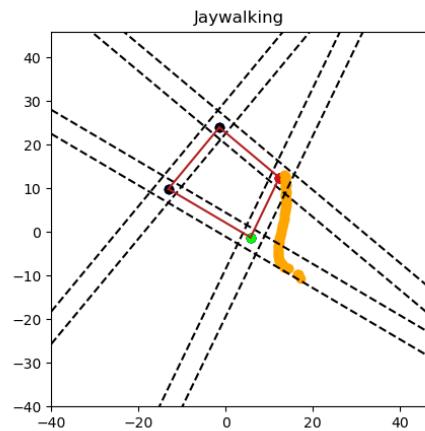
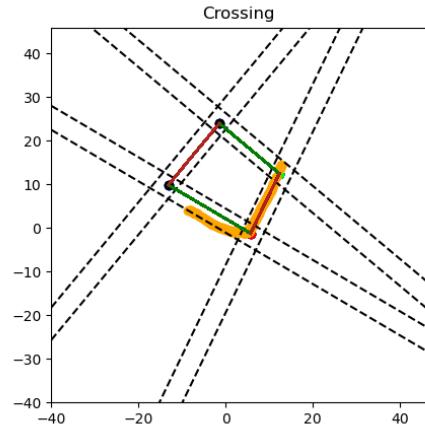


Figure 12. With information on the location of street corners and associated crosswalks, rudimentary rule-based classifiers can distinguish between trajectories of classes such as standard crossing behavior, jaywalking, and bikes passing through on the road. Understanding pedestrian dynamics is crucial to trajectory estimation for safe autonomous vehicle control.

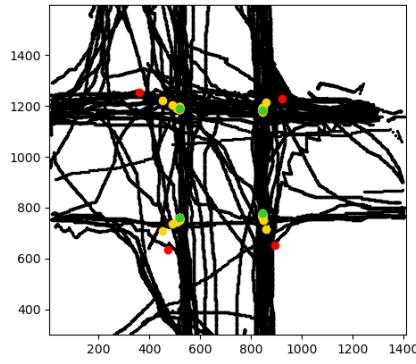


Figure 13. Estimation of crossing corners associated with crosswalks for the traffic circle in the Stanford Drone Dataset, following the same color scheme as described in Figure 8.

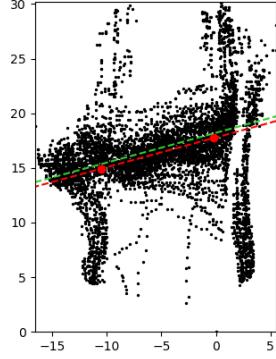


Figure 14. Model representation of the La Jolla crosswalk. Initial estimate using k-means in red, with the final line following iterative refinement shown in green.

rithm's effectiveness at locating the two crossing corners at the La Jolla crosswalk. The initial estimate reflects the noise of those walking on the pathway (notice the density in the lower half of the scatter plot in Figure 14), and after refining the estimate by narrowing the accepted margin of inliers, the crossing better reflects the tendency of pedestrians to gather and orient their trajectory diagonally upward within our bird's-eye-view. Notably, this is not the perpendicular crossing geometry marked by the road, but rather a de facto crossing pattern.

5. Concluding Remarks

In this study, we described and illustrated the effectiveness of an EM algorithm for estimating crosswalks using a real-world pedestrian detection dataset. This algorithm can be used to model both marked and unmarked crossings in a vari-

ety of geometries. Crosswalk detection methods which rely on the physical appearance of road infrastructure naturally fail in cases where the infrastructure is damaged, obstructed, or ignored; by contrast, methods which make use of pedestrian motion can still provide reliable information under these conditions.

5.1. Limitations

While the techniques introduced in the two-stage EM algorithm are designed to make the algorithm robust to certain violations of the prerequisite assumptions, the algorithm is less effective in situations where pedestrians congregate at points within the considered area but away from the crossing origins or destinations. In such cases, the k-means clustering provides a poor initialization (for example, if there are two congregating points at one corner, and a severe imbalance of data at another corner, then two cluster centers will share the same corner and may not converge to the crossing geometry). It would therefore be important to add appropriate check conditions to be sure initial corner estimates are sufficiently spaced. Second, though the best-representative crossing line is determined, further investigation to estimate the effective crossing margin would be valuable, as this is currently passed to the algorithm as a hyperparameter and may vary in different crossing scenes.

5.2. Broader Impact

The demonstrated datasets were collected from both city infrastructure LiDAR and vehicle-mounted camera; any modality which can perform 3D detection is suitable. This method readily extends to measurements made by moving vehicles with accurate localization, and can be used downstream to support safe interactions between pedestrians and autonomous vehicles. As explained in (Gandhi & Trivedi, 2008a), detection and tracking of objects from cameras is easier due to absence of camera motion and use of background subtraction, but in addition to infrastructure-based sensing, improvements in vehicle-to-vehicle and vehicle-to-infrastructure communication will allow for improved exchange of temporal information or information about objects that are seen by one but not seen by others. In particular, vehicles which are able to report pedestrian detections within a global coordinate frame can supply local data to HD maps, leveraging the above algorithm to maintain an accurate, dynamic estimate of de facto crosswalks with repeated samples over time.

6. Acknowledgements

The authors thank the many sponsors and creators of the TRANSFOR 22 dataset, LISA colleagues and Dr. Nachiket Deo for the contribution of pedestrian behavior datasets, and to our reviewers for their constructive comments.

References

- Berriel, R. F., Rossi, F. S., de Souza, A. F., and Oliveira-Santos, T. Automatic large-scale data acquisition via crowdsourcing for crosswalk classification: A deep learning approach. *Computers & Graphics*, 68:32–42, 2017.
- Chen, H. and Zhang, X. Path planning for intelligent vehicle collision avoidance of dynamic pedestrian using att-lstm, msfm, and mpc at unsignalized crosswalk. *IEEE Transactions on Industrial Electronics*, 69(4):4285–4295, 2022. doi: 10.1109/TIE.2021.3073301.
- Coughlan, J. and Shen, H. A fast algorithm for finding crosswalks using figure-ground segmentation. In *2nd Workshop on Applications of Computer Vision, in conjunction with ECCV*, volume 5, 2006.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–22, 1977.
- Deo, N. and Trivedi, M. M. Learning and predicting on-road pedestrian behavior around vehicles. In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1–6, 2017. doi: 10.1109/ITSC.2017.8317865.
- Fridovich-Keil, D., Bajcsy, A., Fisac, J. F., Herbert, S. L., Wang, S., Dragan, A. D., and Tomlin, C. J. Confidence-aware motion prediction for real-time collision avoidance1. *The International Journal of Robotics Research*, 39(2-3):250–265, 2020.
- Gandhi, T. and Trivedi, M. M. Pedestrian protection systems: Issues, survey, and challenges. *IEEE Transactions on Intelligent Transportation Systems*, 8(3):413–430, 2007. doi: 10.1109/TITS.2007.903444.
- Gandhi, T. and Trivedi, M. M. Computer vision and machine learning for enhancing pedestrian safety. In *Computational intelligence in automotive applications*, pp. 59–77. Springer, 2008a.
- Gandhi, T. and Trivedi, M. M. Image based estimation of pedestrian orientation for improving path prediction. In *2008 IEEE Intelligent Vehicles Symposium*, pp. 506–511. IEEE, 2008b.
- Hernández, D. C., Filonenko, A., Seo, D., and Jo, K.-H. Crosswalk detection based on laser scanning from moving vehicle. In *2015 IEEE 13th International Conference on Industrial Informatics (INDIN)*, pp. 1515–1519, 2015. doi: 10.1109/INDIN.2015.7281958.
- Ishak, S., Vlahogianni, E., et al. Trb aed50 data challenge. *TRB Data Challenge*, 2021.
- Ito, T., Tohriyama, K., and Kamata, M. Detection of damaged stop lines on public roads by focusing on piece distribution of paired edges. *International journal of intelligent transportation systems research*, 19(1):56–70, 2021.
- Jo, K., Chu, K., and Sunwoo, M. Gps-bias correction for precise localization of autonomous vehicles. In *2013 IEEE Intelligent Vehicles Symposium (IV)*, pp. 636–641, 2013. doi: 10.1109/IVS.2013.6629538.
- Kitani, K. M., Ziebart, B. D., Bagnell, J. A., and Hebert, M. Activity forecasting. In *European conference on computer vision*, pp. 201–214. Springer, 2012.
- Likas, A., Vlassis, N., and Verbeek, J. J. The global k-means clustering algorithm. *Pattern recognition*, 36(2):451–461, 2003.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- Quintero Mnguez, R., Parra Alonso, I., Fernández-Llorca, D., and Sotelo, M. Pedestrian path, pose, and intention prediction through gaussian process dynamical models and pedestrian activity recognition. *IEEE Transactions on Intelligent Transportation Systems*, 20(5):1803–1814, 2019. doi: 10.1109/TITS.2018.2836305.
- Reisman, P., Mano, O., Avidan, S., and Shashua, A. Crowd detection in video sequences. In *IEEE Intelligent Vehicles Symposium, 2004*, pp. 66–71. IEEE, 2004.
- Ridel, D., Rehder, E., Lauer, M., Stiller, C., and Wolf, D. A literature review on the prediction of pedestrian behavior in urban scenarios. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pp. 3105–3112, 2018. doi: 10.1109/ITSC.2018.8569415.
- Riveiro, B., González-Jorge, H., Martínez-Sánchez, J., Díaz-Vilariño, L., and Arias, P. Automatic detection of zebra crossings from mobile lidar data. *Optics & Laser Technology*, 70:63–70, 2015.
- Robert, L., Yang, J., Tilbury, D., et al. Automated vehicle behavior design for pedestrian interactions at unsignalized crosswalks. Available at SSRN 3859366, 2021.
- Robicquet, A., Sadeghian, A., Alahi, A., and Savarese, S. Learning social etiquette: Human trajectory understanding in crowded scenes. In *European conference on computer vision*, pp. 549–565. Springer, 2016.

- Satzoda, R. K. and Trivedi, M. M. Drive analysis using vehicle dynamics and vision-based lane semantics. *IEEE Transactions on Intelligent Transportation Systems*, 16(1):9–18, 2014.
- Sivaraman, S. and Trivedi, M. M. Dynamic probabilistic drivability maps for lane change and merge driver assistance. *IEEE Transactions on Intelligent Transportation Systems*, 15(5):2063–2073, 2014.
- Song, H., Jo, K., Cho, J., Son, Y., Kim, C., and Han, K. A training dataset for semantic segmentation of urban point cloud map for intelligent vehicles. *ISPRS Journal of Photogrammetry and Remote Sensing*, 187:159–170, 2022. ISSN 0924-2716. doi: <https://doi.org/10.1016/j.isprsjprs.2022.02.007>. URL <https://www.sciencedirect.com/science/article/pii/S0924271622000417>.
- Suzuki, S., Raksincharoensak, P., Shimizu, I., Nagai, M., and Adomat, R. Sensor fusion-based pedestrian collision warning system with crosswalk detection. In *2010 IEEE Intelligent Vehicles Symposium*, pp. 355–360, 2010. doi: 10.1109/IVS.2010.5548120.
- Tian, R., Li, L., Yang, K., Chien, S., Chen, Y., and Sherony, R. Estimation of the vehicle-pedestrian encounter/conflict risk on the road based on tasi 110-car naturalistic driving data collection. In *2014 ieee intelligent vehicles symposium proceedings*, pp. 623–629. IEEE, 2014.
- Tian, S., Zheng, M., Zou, W., Li, X., and Zhang, L. Dynamic crosswalk scene understanding for the visually impaired. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29:1478–1486, 2021. doi: 10.1109/TNSRE.2021.3096379.
- Tmen, V. and Ergen, B. Intersections and crosswalk detection using deep learning and image processing techniques. *Physica A: Statistical Mechanics and its Applications*, 543:123510, 2020. ISSN 0378-4371. doi: <https://doi.org/10.1016/j.physa.2019.123510>. URL <https://www.sciencedirect.com/science/article/pii/S0378437119319582>.
- Zang, G., Azouigui, S., Saudrais, S., Peyricot, O., and Hébert, M. Quantitative study of vehicle-pedestrian interactions: Towards pedestrian-adapted lighting communication functions for autonomous vehicles. *Electronic Imaging*, 2021(17):172–1, 2021.
- Zhai, Y., Cui, G., Gu, Q., and Kong, L. Crosswalk detection based on mser and eransac. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, pp. 2770–2775, 2015. doi: 10.1109/ITSC.2015.448.
- Zhu, D., Zahran, M., Li, L. E., and Elhoseiny, M. Motion forecasting with unlikelihood training in continuous space. In Faust, A., Hsu, D., and Neumann, G. (eds.), *Proceedings of the 5th Conference on Robot Learning*, volume 164 of *Proceedings of Machine Learning Research*, pp. 1003–1012. PMLR, 08–11 Nov 2022. URL <https://proceedings.mlr.press/v164/zhu22a.html>.
- Zhu, H., Iryo-Asano, M., Alhajyaseen, W. K., Nakamura, H., and Dias, C. Interactions between autonomous vehicles and pedestrians at unsignalized mid-block crosswalks considering occlusions by opposing vehicles. *Accident Analysis & Prevention*, 163:106468, 2021.