TRB 100th Annual Meeting (January 25-29, 2021)

Smart Transportation System Laboratory
Tel: +82-42-350-3674
Department of Civil & Environmental Engineering
Korea Advanced Institute of Science and Technology

Analysis of Vehicle-Pedestrian Interactive Behaviors near Unsignalized Crosswalk

Byeongjoon Noh
powernoh@kaist.ac.kr

David Lee
david733@alum.mit.edu

Dongho Ka
kdh910121@kaist.ac.kr

Hwasoo Yeo*
hwasoo@kaist.ac.kr

Introduction

❑ Observation

◆ High pedestrian casualties from traffic accidents

- Death or injuries over 50 million people each year from traffic accidents [1]
- Especially, pedestrians are exposed to the risk of death on the roads
  - ➢ Drivers failing to yield to pedestrians at the crosswalks
  - ➢ Blind spots near the roadsides, etc.
- In Korea, pedestrian accident is the highest percentage of accident type [2]

[1] Ho GT, Tsang YP, Wu CH, Wong WH, Choy KL. A computer vision-based roadside occupation surveillance system for intelligent transport in smart cities. Sensors. 2019 Jan;19(8):1796.

[2] Lytras MD, Visvizi A. Who uses smart city services and what to make of it: Toward interdisciplinary smart cities research. Sustainability. 2018 Jun;10(6):1998.

◆ Various ways to reduce and protect pedestrian accidents

- Analyzing accident historical data
  - ➢ Deploying bumpers and speed cameras which suppress engaging in risky or illegal behaviors of drivers and pedestrians
- Vision-based analytical system
  - ➢ Operating 24-hour closed-circuit television (CCTV) surveillance center at districts

❑ However

◆ Post-factor perspectives

- Used only historical data of traffic accidents to reinforce safe urban environment
  - ➢ Deploying fences, additional CCTVs, etc.
- ➔ Need to "proactive strategies"

◆ High cost and time to handle video

- Difficult to precisely extract objects' behavioral features in oblique views
  - ➢ *e.g.*, vehicle velocities, pedestrian positions, etc.
- Relied on manual inspection to extract them from video footage
  - ➢ Required much more cost and time when expended to the urban scale
- ➔ Need to "automated feature extraction"

❑ Purpose

◆ New analytical system for pedestrian's potential risk using vision

- To automatically extract the traffic-related objects' behavioral features
  - ➢ Affecting the likelihood of potentially dangerous situations after detecting them
- To analyze their behavioral features and relationships among them by locations
- To support efficient decision-making to improve road environments safer
  - ➢ By providing the object movement patterns in unsignalized crosswalk

❑ Contributions

◆ Implementation of automated feature extraction system

- With computer vision and deep learning techniques
- From "motioned scenes" selection and feature extraction

◆ Analysis of potential risks in actual camera locations

- Covering five cameras over two weeks in Osan city, South Korea
- Compared the behaviors of vehicles and pedestrians by visualizing the changes of behaviors

◆ A proactive pedestrian safety system

- By deriving the useful information for pedestrian risks
  - ➢ Can provide decision makers with it, and make safer road environment
- First step to reach in an advanced transportation safety system

Materials and Methods

❑ Overall system architecture

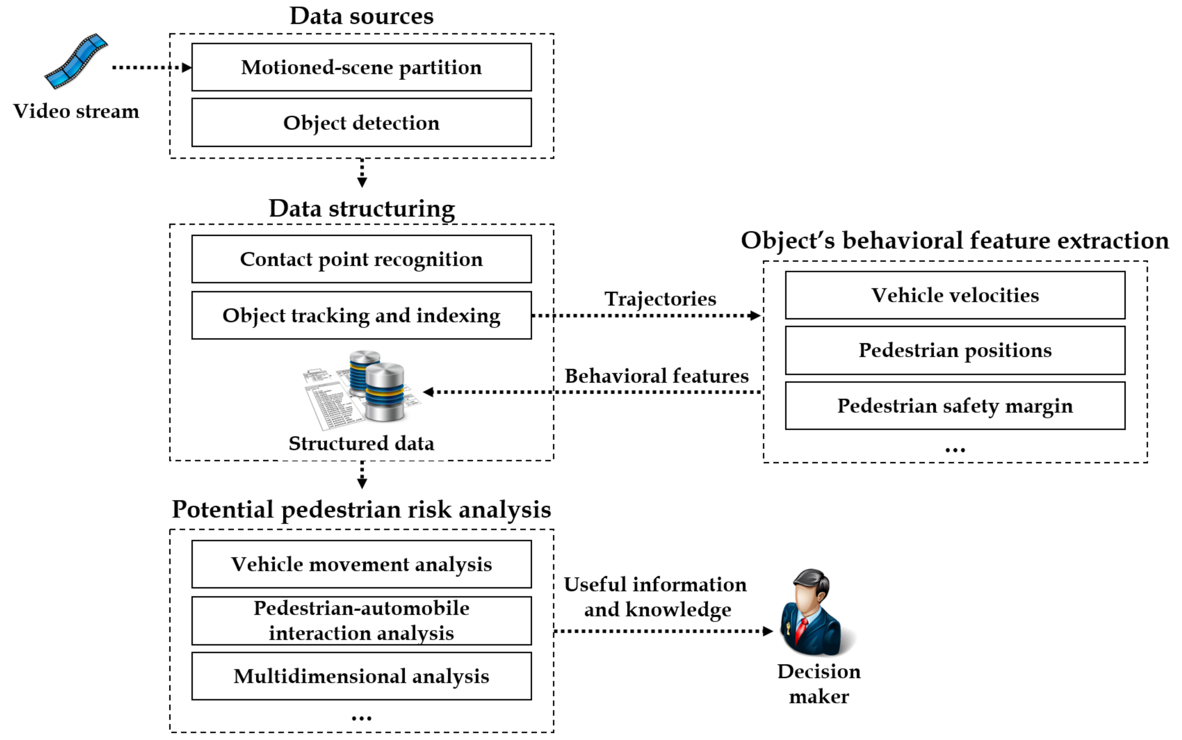
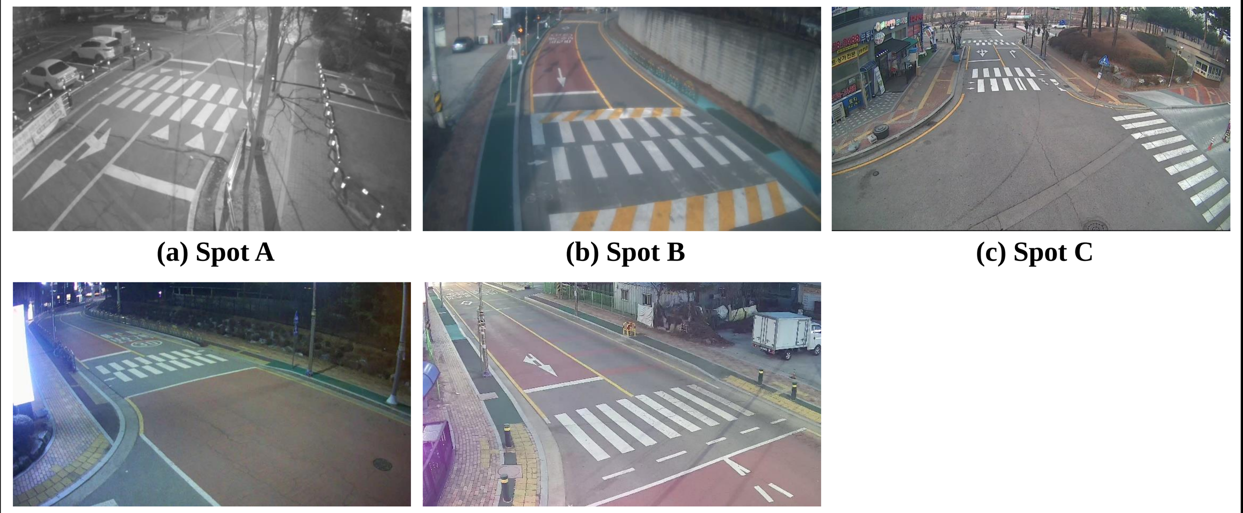


Figure 1. Overall architecture of the proposed system

❑ Data sources

◆ Target spots and their environmental characteristics

- Used video data from CCTV cameras over five unsignalized crosswalks
  - ➢ In Osan city, South Korea (Spots A to E)
  - ➢ Intended to record and deter instances of street crime
  - ➢ Collected video from January 9th to January 28th during commuting hours



(a) Spot A (b) Spot B (c) Spot C (d) Spot D (e) Spot E

Figure 2. Actual CCTV views in (a) Spot A to (e) Spot E

Spot code	Crosswalk width (m)	School zone	Speed cam.	Fence	Red urethane	Raised crosswalk	# of lanes	Speed limits (km/h)
A	$\approx 15m$	√	X	√	X	√	2	
B	$\approx 15m$	√	X	√	√	√	2	
C	$\approx 15m$	X	X	X	X	X	2	$30\ km/h$
D	$\approx 15m$	√	X	√	√	X	2	
E	$\approx 20m$	√	X	√	√	X	2	

Table 1. Information of characteristics of spots

◆ Scene clips extraction

- Extracted only view clips ("scenes") with moving vehicle activity from video stream
  - ➢ Due to handling huge volume of video data recorded in multiple areas for long time periods
- Frame difference: $|P[I(t)] - P[I(t+1)]| > Threshold$
  - ➢ where image obtained at the time t denoted by $I(t)$, $P[*]$ is pixel value in image
  - ➢ A simple & low computational complexity
  - ➢ Commonly used approach for recognizing the movements in video [3], [4]

Figure 3. Example of frame difference between foreground and background

[3] Liu H, Dai J, Wang R, Zheng H, Zheng B. Combining background subtraction and three-frame difference to detect moving object from underwater video. InOCEANS 2016-Shanghai 2016 Apr 10 (pp. 1-5). IEEE.

[4] Sengar SS, Mukhopadhyay S. Moving object detection based on frame difference and W4. Signal, Image and Video Processing. 2017 Oct 1;11(7):1357-64.

◆ Object detection and segmentation using deep learning

- Detected the traffic-related objects (e.g., vehicles and pedestrians in this paper)
  - ➢ With the motioned video clips using frame difference
  - ➢ Used a mask R-CNN model in Detectron 2 platform provided by Facebook AI research (FAIR) [5]

[5] Facebook AI Research. Available from: https://ai.facebook.com/ Accessed Apr 12, 2020.

Spot code	# of the extracted scenes	
	Vehicle-only scenes	Interactive scenes
A	4,632	2,322
B	2,481	1,395
C	3,533	2,079
D	1,843	1,002
E	4,572	3,203

Table 2. The numbers of the extracted scenes

❑ Data structuring

◆ Contact point recognition [6]

- Contact point: reference points for determining their features
- To extract the precise behavioral features
- Obtained by perspective-transforming ground tip points into an overhead view
  - ➢ Ground tip: a point on the ground directly underneath the center of the object in the oblique view
  - ➢ Transformation matrix is used in OpenCV library with four anchor points

◆ Object tracking and indexing [6]

- By using Kalman filter-based tracking algorithm
- Step 1: Estimating the candidate points based on smoothing
- Step 2: Assigning objects in next frame by calculating and comparing distances

[6] Noh B, No W, Lee J, Lee D. Vision-Based Potential Pedestrian Risk Analysis on Unsignalized Crosswalk Using Data Mining Techniques. Applied Sciences. 2020 Jan;10(3):1057.

❑ Automated object behavioral feature extraction

◆ Used features and their extraction methods

- Vehicle velocity
  - ➢ $v_i^{j,(j+1)} = {}^{D_{pixel}(point_i^j, point_i^{(j+1)})}/_{F*P}$ (m/s)
- Pedestrian position
  - ➢ Categorized "sidewalk", "crosswalk", "crosswalk influenced area (CIA)", and "road" (see Figure 5)
- Vehicle–pedestrian and –crosswalk distance
  - ➢ $Dist_{i,k}^j = {}^{D_{pixel}(object_i^j, object_k^j)}/_{P}$ (m)

$D_{pixel}(A, B)$: Euclidean distance between A and B
F: time conversion constant
P: pixel-per-meter constant

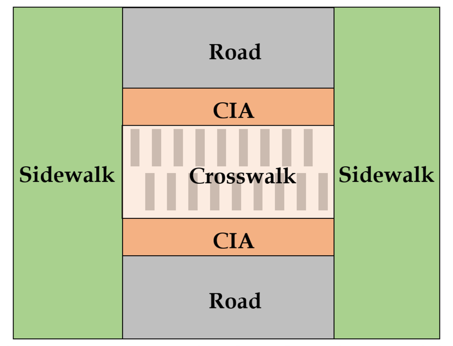


Figure 5. Categorized pedestrian position

Experiments and Results

❑ Experimental design

◆ Basic analysis toward to vehicle velocities

- Separated the average vehicle velocities into vehicle-only and interactive scenes
  - ➢ As described in Figure 6
- In all spots, velocities in interactive scenes were lower than those in car-only scenes
- In most spots, vehicles traveled under the speed limits (30km/h) except for Spot E
- In Spot C, vehicles moved slowly ➔ As we guess, drivers may concern that pedestrians suddenly jump out the road

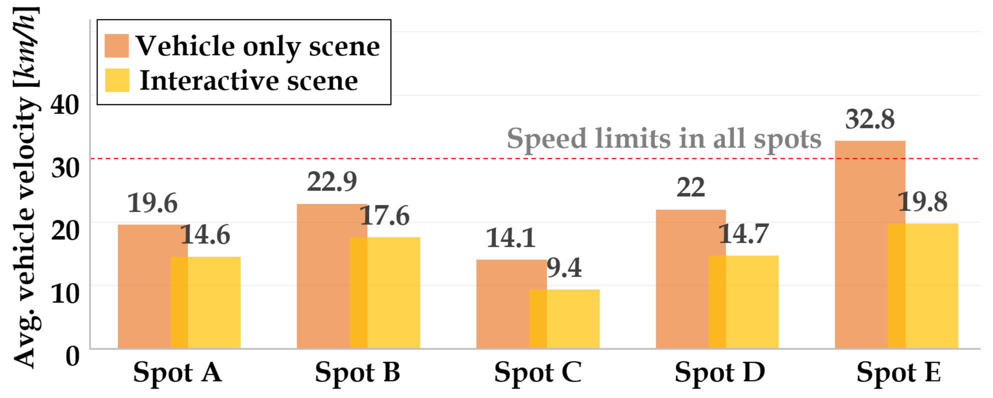


Figure 6. Average vehicle velocities in each spot by scene type

◆ Analysis scenario

- Scenario I: Analyzing changes in vehicle velocities near unsignalized crosswalks
  - ➢ Focused on analyzing vehicle velocity changes near the crosswalk in each spot when there are no pedestrians present (vehicle-only scenes)
- Scenario II: Analyzing vehicle velocity changes by vehicle-pedestrian distances near unsignalized crosswalks
  - ➢ To investigate how vehicles behaved when pedestrians were on the crosswalks

❑ Results

◆ Analysis of scenario I

- Measured the velocities of the vehicles at certain distances from crosswalks (vehicle-crosswalk distance) Because each scene has a different number of frames
  - ➢ Baselines: 10*m*, 5*m*, 3*m*, 0*m* (starting line of the crosswalk)
- Observed velocity changes before, during, and after the vehicles pass the crosswalk in each area (see Figure 7)

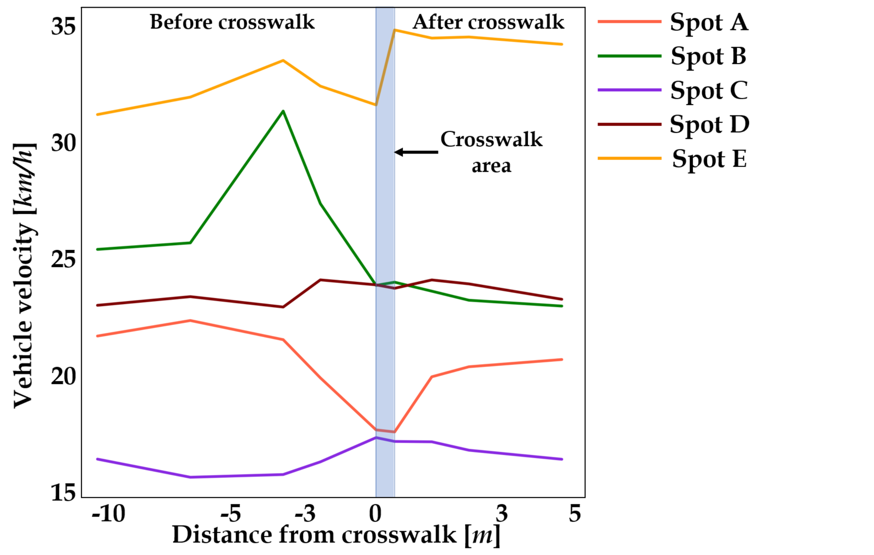
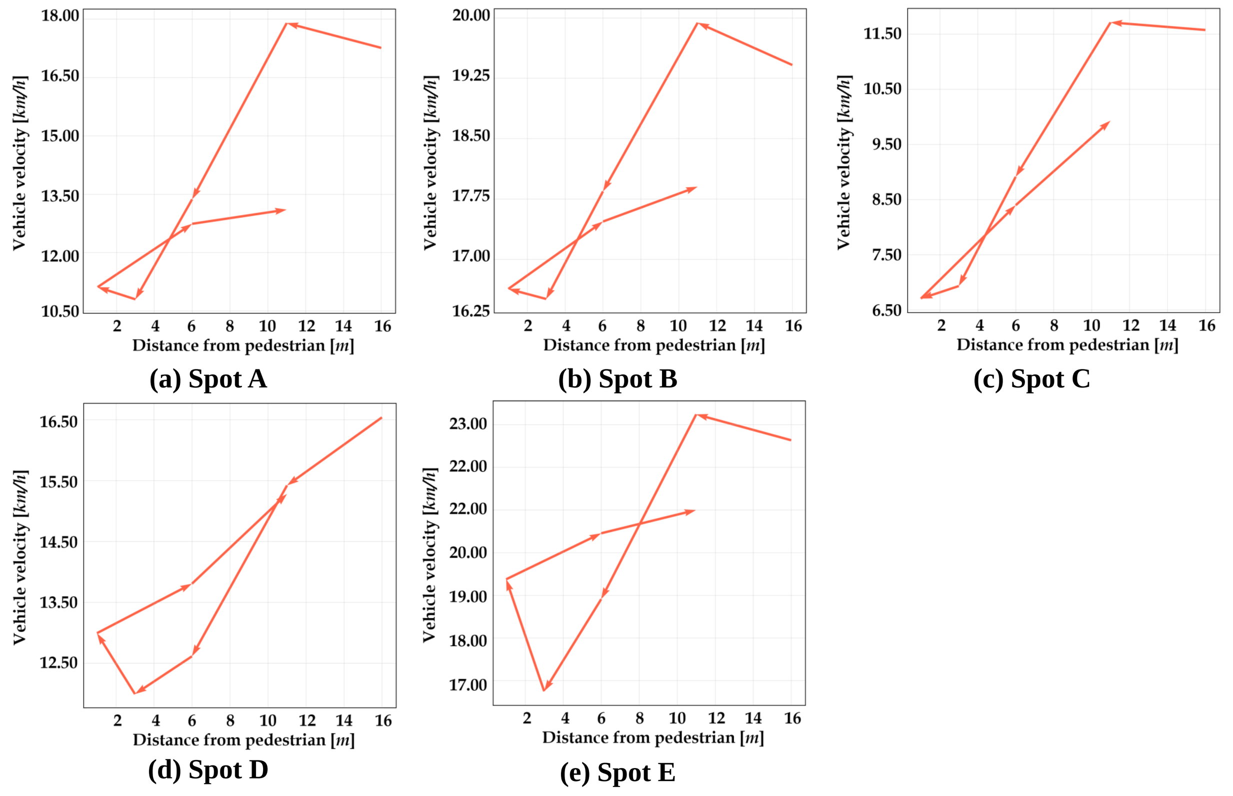


Figure 7. Results of average vehicle velocity changes by distances from crosswalk in each spot

- Average vehicle velocity never exceeded the speed limit except for Spot E
- Vehicles in spots A and B slowed down faster before crosswalks
  - ➢ Still decreased while passing the crosswalks whereas Spot E accelerated again
  - ➢ ➔ Hypothesized that the raised crosswalks were forcing drivers to alleviate the impact by rapidly decelerating (see Figure 2 (a) and (b))
- In Spot C, vehicles decelerated at a distance of about 3-5m rather than just in front of crosswalk, thereafter increases again
  - ➢ Can be interpreted that when the vehicles approach the crosswalk, they slowed down in advance and accelerate again after checking that there is no pedestrian

◆ Analysis of scenario II

- Targeted the scenes at unsignalized crosswalks when pedestrians were in the crosswalks
- Extracted the vehicle velocities from frames when the vehicle-pedestrian distances were between 2m to 16*m*
- Observed a "lasso-like" pattern in all spots (See Figure 7)
  - ➢ Vehicles dramatically slowed down when approaching a pedestrian, then increased their speeds just before reaching their closest distance
- Need a facilities that can protect pedestrians near crosswalks!



(a) Spot A (b) Spot B (c) Spot C (d) Spot D (e) Spot E

Figure 7. Results of average velocity changes by distances between vehicles and pedestrians in (a) Spot A to (e) Spot E, respectively

Conclusions

❑ Summary

◆ A new designed potential risk analytical system

- Can support to make efficient decisions in order to improve road environment safer
- Based on vision

◆ Main objectives

- To automatically extract the traffic-related objects' behavioral features
  - ➢ Affecting the likelihood of the potentially dangerous situations
- To analyze their behavioral features and relationships among them
- To support the administrators make efficient decisions to improve the road environment safer

◆ Results

- Visualized and compared changes of vehicle velocities near unsignalized crosswalks in each camera locations
- Obtained the patterns of vehicle movements according to road environment and pedestrian presence

❑ Expected results

◆ Automated video handling system

- With computer vision and deep learning techniques
- Validated feasibility and applicability in our experiments

◆ Support decision maker to make efficient decisions

- Can provide traffic engineers and urban designers with clues to help further investigation using data analysis techniques
- With broad and dense analysis of behaviors as a proactive pedestrian safety system