Collision prediction with oncoming pedestrians on Braille blocks

Tomoya Ishii 1, Hidehiko Shishido 2,3 and Yoshinari Kameda 2

1 Master’s program in Intelligent and Mechanical Interaction Systems, University of Tsukuba, 1-1-1 Tennoudai, Tsukuba, Ibaraki, 305-8573, Japan
2 Center for Computational Sciences, University of Tsukuba, 1-1-1 Tennoudai, Tsukuba, Ibaraki, 305-8573, Japan
3 Currently at Soka University

Abstract
This research targets blind people walking on braille blocks. Some non-blind pedestrians walk on braille blocks while they watch their smartphones on the street. Blind people may collide with these oncoming pedestrians. When the oncoming pedestrian does not notice the blind people, a collision will likely occur. We propose a new method for predicting collisions in such situations. We use a smartphone's camera to predict collisions. To predict the collision, we set two conditions. The first condition is whether the oncoming pedestrian is on the collision path. The second condition is whether the oncoming pedestrian notices the blind people.

Keywords
Collision prediction, distance estimation, path estimation, gaze estimation

1. Introduction
This research targets blind people walking on braille blocks. Blind people rely on braille blocks when they go out. Some non-blind pedestrians walk on braille blocks while they watch their smartphones on the street. It is difficult for blind people to avoid these pedestrians and may collide with them. In a potential collision situation, oncoming non-blind pedestrians should give way. Braille blocks are installed to help blind people walk safely in Japan [1, 2]. When oncoming pedestrians notice blind people, they are asked to give their way to avoid collisions.

We propose a new method for predicting collisions with oncoming pedestrians using a smartphone’s camera. To predict collision, we set two conditions. The first condition is whether the oncoming pedestrian is on the collision path. The second condition is whether the oncoming pedestrian notices the blind people. We foresee that a collision will occur when the oncoming pedestrian is on the collision path and does not notice the blind people.

Once the collisions can be predicted, a loud sound signal can warn both people. The warning can make spare time for blind people to protect themselves in case of a collision. The warning can also ask the oncoming pedestrian to avoid a collision. As the loud sound signal causes a big stress on all the people on the street, the collision prediction should be accurate, and the collision prediction system calls the warning at the last moment when the collision is inevitable.

2. Related work
2.1. Collision avoidance for blind people

Various types of obstacle avoidance systems for blind people have been proposed. These include the systems to detect obstacles by attaching an ultrasonic sensor [3, 4] or LiDAR [5]...
Collision avoidance systems using a suitcase-shaped device have been proposed [6-8]. The suitcase has a stereo camera, LiDAR [6, 8], and a laptop computer. The BBEEP system [7] predicts collision by estimating the future location of oncoming pedestrians based on their walking trajectories. The system beeps and asks pedestrians to give their way to avoid a collision. Vibrators [6] and levers [8] are equipped on the suitcase handles to indicate the path for blind people. These tactile feedbacks enable them to avoid a collision on their own. These are promising approaches in case the equipment can accompany blind people.

Collision avoidance systems using smartphones [9, 10] and wearable devices [11] have been proposed. These systems use LiDAR on smartphones and stereo cameras equipped to the chest of blind people to measure the distance to the object. These systems focus on obstacles and pedestrians at a short distance and do not consider oncoming pedestrians walking from a distance.

### 2.2. Collision avoidance in autonomous mobile robots

As for obstacle collision avoidance, systems on autonomous mobile robots have been proposed. Ultrasound sensors [12, 13] and LiDAR [14, 15] detect obstacles. The systems are large because they use specific sensors to measure the distance of distant obstacles. It is challenging to adopt these methods directly for humans.

Collision avoidance with dynamic obstacles [16] and moving pedestrians [17] have been proposed too. The robot is set to go away from the original planned path or make a curve to avoid collisions. We should not instruct blind people to change their paths because they may lose their way once they are off the braille blocks.

### 2.3. Gaze estimation

The gaze plays an important role in human interaction. People select collision-free paths by considering the gaze of other pedestrians and the direction they walk [18-20].

Many studies have been proposed on appearance-based gaze estimation using deep learning [21-28]. Appearance-based gaze estimation requires datasets containing a variety of environments, subjects, and targets [21, 23, 25, 27, 28]. Zhang et al. provided a gaze image dataset of participants acquired while they watched a laptop computer in their daily life and proposed a gaze estimation method [21]. Sugano et al. estimated gaze on a public display without individual user calibration [22]. Recasens et al. estimated which objects the user looked at in the image [23]. Chong et al. extended it to video and correctly detected the gaze target even outside the image [24].

Kellnhofer et al. provided a dataset captured at a wide range of head postures and distances to achieve 3D gaze estimation [25], where the eyes may not be visible by cameras such as surveillance cameras due to occlusion. 3D gaze estimation [26] and target object estimation [27] have been proposed in situations where only the back of the head is visible. Bermejo et al. approximated the gaze by the posture of the head [26]. Nonaka et al. estimated the gaze by considering the body orientation [28].

### 3. Collision avoidance on braille blocks

Braille blocks are installed in a straight line [2]. Once the blind people become on braille blocks, they follow the blocks. When oncoming pedestrians notice blind people, they should give their way to avoid collisions because the block is installed to support the safe walking of blind people. As shown in Figure 1, a collision occurs when an oncoming pedestrian walks on braille blocks without noticing the blind people.

In our proposal, blind people wear their smartphones at chest height, as shown in Figure 2. The smartphone’s camera is tilted downward from the horizontal. This tilt allows the camera to capture the walking area in front of the user. The setup of the smartphone in this way does not interfere with their walking style.

The proposed system uses only a smartphone to cover the process from video acquisition to collision prediction.
4. Oncoming pedestrians on the collision path

4.1. Detection

Pedestrian detection methods from various camera viewpoints have been proposed [29-33]. Pedestrians are detected for tracking [29, 30], counting [31], and autonomous driving systems [32]. Static obstacle detection methods have been proposed from a pedestrian’s viewpoint [34, 35]. We can apply such methods to static objects. In this research, we focus on the detection of oncoming pedestrians.

A detector with high-speed detection is required to achieve real-time detection. We need a new detector that finds pedestrians on braille blocks. Note that the braille blocks are square.

The system should first find braille blocks. We utilize YOLOv7 [36] as it can run fast. We use the braille block dataset [37] to train YOLOv7. We split 2000 images 4:1 for training and validation. The batch size is 16. The number of epochs is 150. The image size was resized from 1024 x 1024 to 512 x 512 for training.

Once the trained YOLOv7 find more than two braille blocks, the braille block region can be found. The region is defined by two sidelines of the braille blocks as the blocks form a straight line on the street. The left and right vertices of the bottom edge of the detected blocks are used to estimate the sidelines of the braille block region. The sidelines are estimated by the least-squares method. The region is the collision path.

We also use YOLOv7 to detect pedestrians. The center of the bottom edge of the detected pedestrian rectangle is counted as the pedestrian's foot position. If the foot position is inside the braille block region, we detect the pedestrian as an oncoming pedestrian on the collision path.

In case of finding less than two braille blocks, the braille block region found in the last frame is used to detect the oncoming pedestrian on the collision path.

Figure 3 shows the detection results of oncoming pedestrians. The light blue bounding boxes indicate the braille blocks detected by the trained YOLOv7. Note that the bottom edge of the braille block rectangle is marked by blue color. The red crossing lines indicate the sidelines of the braille block region, inside of which is the collision path.

Figure 3 (a) shows a situation where an oncoming pedestrian is on the collision path. The person wearing black is on the collision path and is detected. It is marked with a red bounding box. As the oncoming pedestrian approaches, we can detect the smartphone they hold, as indicated by the green bounding box by the YOLOv7.

Figure 3 (b) shows a situation where the oncoming pedestrian is not on the collision path. The person wearing white is not on the collision path and is marked by the blue bounding box.

Oncoming pedestrians appear larger in the image as they approach. Oncoming pedestrians at close locations likely hide most of the braille blocks. In such a case, the braille block region detected in the previous frame will be used.

Braille blocks may not be installed in a straight line. Based on the guideline [2], braille blocks are rarely curved in Japan. Therefore, we assume the sidelines of the braille block region can be approximated almost as straight lines.
4.2. Distance estimation

We need to estimate the distance to oncoming pedestrians to predict a collision with oncoming pedestrians. Methods for estimating the distance to an obstacle using a camera have been proposed [38-40]. Detected facial features [38] and rectangle size [39, 40] are used. Chen et al. used the camera’s focal length, angle of view, and information on its orientation relative to the ground [34]. Combined with the assumption that the road surface is horizontal, the distance to the obstacle can be estimated.

In this research, we use the method [34] and recognize an oncoming pedestrian as an obstacle. We estimate the distance from the blind people to the position where the oncoming pedestrian stands, as shown in Figure 4.

4.3. Future position estimation

Blind people walk on braille blocks installed in a straight line. If oncoming pedestrians also walk on braille blocks, they walk straight toward blind people. We must decide whether the oncoming pedestrian walks straight toward the blind people. We estimate the future position of the oncoming pedestrian. The distance of the collision with the oncoming pedestrian is about 60 cm, which is within the reach of a white cane. Therefore, the distance and future position estimations must be accurate enough to meet this requirement.

Some research proposed a method for estimating the trajectory of pedestrians from an egocentric video [41, 42]. Yagi et al. estimate the position of a pedestrian's waist using the pedestrian's skeletal information and the camera's pose information [41]. Qiu et al. estimate the future position by referring to the property of a rectangle instead of a point [42]. This method can be combined with the distance estimation method [34] to estimate the walking trajectory.
5. The decision of whether oncoming pedestrians notice blind people

Suppose the oncoming pedestrian is on the collision path. In that situation, the remaining problem is to decide whether the oncoming pedestrian notices the blind people and intends to avoid the collision. Our proposed system estimates the gaze direction of the oncoming pedestrians. It checks whether the gaze direction is toward the location of the blind people.

Lee et al. have proposed a system to decide whether an oncoming pedestrian is looking at blind people [38]. Their system is trained by the annotations of the pedestrian’s face images. The direction of gaze has been estimated as a 3D vector [25, 26, 28, 43]. As shown in Figure 5, Zhang et al. achieve gaze estimation even when the resolution of the cropped face image is low [43].

We plan to adapt the method [43] to gaze estimation of the oncoming pedestrians to decide whether they notice blind people. Gaze estimation should start at the moment when the oncoming pedestrian is detected, up to the time when the collision occurs.

Figure 5: Results of gaze estimation from face images (Figure 8 in [43]).

6. Collision prediction with oncoming pedestrians

To predict collision, we set two conditions. The first condition is whether the oncoming pedestrian is on the collision path. The second condition is whether the oncoming pedestrian notices the blind people. Even when the oncoming pedestrian keeps the collision path, the system does not make a warning once it detects the oncoming pedestrian gaze the blind people just at a frame. The system calls the warning of collision if the oncoming pedestrian comes within the hazardous distance of the blind people without even a glance at the blind people in their front.

7. Conclusion

We proposed a new method for predicting collisions with oncoming pedestrians using a smartphone’s camera. To predict collision, we set two conditions. The first condition is whether the oncoming pedestrian is on the collision path. The second condition is whether the oncoming pedestrian notices the blind people. We developed a system for the first condition and showed the snapshots of the results.

We plan to incorporate the procedure of the second condition into our system. Implementing the total system on a smartphone will help blind people to avoid collisions with oncoming pedestrians.

Part of this research is supported by JSPS Kaken 22K19803.

8. References


Object Counting Invariance of Conv

A. G. Hauptmann

Z. Q. Cheng, Q. Dai, H. Li, J. Song, X. Wu,


[38] K. A. Rahman, M. S. Hossain, M. A. Bhuiyan, T. Zhang, M. Hasanuzzaman, H. Ueno, Person to Camera Distance Measurement Based on Eye-Distance, in:


