Inter-Vehicle Communication Protocol for Cooperatively Capturing and Sharing Intersection Video

Kazuya Kotani, Weihua Sun, Tomoya Kitani[†], Naoki Shibata[††],
Keiichi Yasumoto, and Minoru Ito

Graduate School of Information Science, Nara Institute of Science and Technology
[†]*Division of Global Research Leaders, Shizuoka University*
[††]*Department of Information Processing and Management, Shiga University*

Abstract

For accident prevention at intersections, it is useful for drivers to grasp the position of vehicles in blind spots. This can be achieved without infrastructure if some vehicles passing near the intersection capture and share live video of the intersection through inter-vehicle communications. However, such video streaming requires a congestion control mechanism. In this paper, aiming to let a driver grasp the situation at an intersection, we propose a method to select vehicles that send a video in order to generate a live bird's-eye-view video of the intersection. In our method, each vehicle at an intersection exchanges information with others, such as the sub-areas of the intersection it captures, the quality of its video, and its position and speed. Based on the exchanged information, each vehicle autonomously judges whether it should send its video or not. Through simulation with a QualNet simulator, we confirm that our method achieves a good video arrival rate and video quality sufficient for practical use.

1 Introduction

Most traffic accidents occur at intersections, and the ratio of near-intersection accidents to all traffic accidents is about 40%. The main cause of near-intersection accidents is that there are vehicles, motorcycles, bicycles and pedestrians hidden in blind spots that go unnoticed by drivers. Hereafter, we call these vehicles in blind spots *"Vehicles In Blind Spot" (VIBS)*. In the field of Intelligent Transport Systems (ITS)[1], traffic accident prevention by detecting VIBS is a hot topic. Although some vision support methods for safe driving have been proposed, most of these methods focus on providing vision around an individual vehicle by using multiple in-vehicle cameras. In order to provide a useful real-time video of VIBS to drivers, multiple vehicles in the intersection should cooperate to compose videos captured from multiple vehicles and share the video wirelessly.

In this paper, we aim to make drivers aware of the positions of VIBS. We propose a method for composing videos captured from multiple vehicles at an intersection into a bird's-eye-view video and sharing that video wirelessly among neighboring vehicles. In order to make the method available at any intersection without infrastructure, we design the method as a fully distributed system utilizing inter-vehicle communication.

In order to efficiently use the wireless bandwidth to share high quality videos, the most suitable vehicles are selected to capture videos in the proposed method. For this purpose, we divide the intersection into uniform rectangular cells, and vehicles exchange their information with regard to position, orientation, speed and video quality. The only vehicles determine the priority of each vehicle. A vehicle that can capture more cells with higher picture quality will be assigned a higher priority. Then, vehicles with high priorities send their videos to neighboring vehicles.

To evaluate our method, we conducted simulations. We reproduced a realistic environment with obstacles and moving vehicles. The wireless network is simulated using the QualNet simulator [2]. We evaluated the number of cells in the composed video, the average quality of the video, etc. We compared the following three methods of choosing suitable vehicles to capture and send their video: (i) all vehicles in the intersection, (ii) the vehicles closest to the center for four directions and (iii) the proposed method. We confirmed that the proposed method saves bandwidth for sharing a higher quality composed video containing more cells.

2 Related Work

The major car makers have been developing ITS technologies, especially driving safety support systems (DSSS).

Honda Motor Co., Ltd. has been working on an ASV project promoted by the Japanese government since 1991. Some vehicles with DSSS developed by this project were shown as demonstrations in "ITS-Safety2010" [3].

Some car models manufactured by Nissan Motor Co., Ltd. have an "Around View Monitor" [4]. This system provides vision around the car by using multiple in-vehicle cameras. However, there are no studies published for allowing multiple vehicles to share videos captured by themselves via wireless communication.

As a study for safety driving support, Ota et al. proposed a method for composing the intersection pictures captured by in-vehicle cameras into bird's eye view pictures when multiple vehicles enter an intersection [5]. However, they did not mention how to exchange the video data in their paper.

In [6], Meng et al. proposed a video streaming method utilizing inter-vehicle communication. This method streams video with inter-vehicle communication when a vehicle requests a live video from a remote place. This method selects vehicles that forward video based on destination, mobility, etc. of other vehicles to efficiently use wireless bandwidth. However, they did not consider sending video in real time, and this method is designed for sending video on a highway.

Our contribution in this paper is that we propose an inter-vehicle protocol that allows each vehicle to acquire only important videos to compose a bird's-eye-view video from other vehicles within available bandwidth. Our method uses the technique proposed in [5] to create a bird's eye view video. In this paper, we focus on achieving a real-time video by exchanging captured video among vehicles.

3 Method for Cooperative Video Capturing and Sharing at Intersections

In this section, we show the overview of our proposed method, give the assumptions and the formulation of the problem. Then we explain the proposed method for cooperatively capturing and sharing an intersection video.

3.1 Overview

The goal of our study is to prevent traffic accidents in an intersection by providing a bird's-eye-view video to drivers as shown in Fig. 1. As the target situation, we suppose an environment in which many vehicles are waiting to turn at a four-way intersection. For the intersection shown in Fig. 2, our proposed protocol works as follows.

1. Vehicle 1 broadcasts a request message specifying an area covering blind spots before it turns right.
2. Vehicles 2, 3 and 4 are assigned the highest priorities among all vehicles and selected as video sender vehicles. Then, they broadcast the videos they captured.

3. Vehicle 1 receives the videos from Vehicles 2, 3 and 4, and composes these videos into a bird's-eye-view video.

3.2 Assumptions

We assume that the target intersection is divided into uniform rectangular cells and each vehicle knows the cell boundary from the car navigation system (Fig. 3). We also assume that a bird's-eye-view video can be composed from videos recorded from multiple directions by using the video compositing techniques in [5]. In this section, we describe assumptions about the target intersection and vehicles.

Although the measurement error of GPS receivers used in current car navigation systems is usually about 10 [m], we assume that there is no positioning error in this paper, since there are advanced methods for accurate positioning with GPS [7].

The moving speed of each car is assumed to be less than 45 [km/h]. The distance between the moving vehicles is more than 5 [m]. Each vehicle knows the driver's next action (e.g. turning right, going straight, etc.) when approaching an intersection from the turn-signal switch, the scheduled route in the car navigation system, etc.

We assume that the set of cells captured by an in-vehicle camera can be calculated from the view angle of the camera, the vehicle's position and orientation, and positions of the other vehicles in the intersection.

We assume that each vehicle has the following equipments.

- In-vehicle camera with QVGA resolution
- Car navigation system with GPS, a digital map, and a display
- IEEE 802.11 WLAN communication device
- In-vehicle computer with sufficient storage

3.3 Problem Formulation

In order to achieve the proposed method, we solve the following problem: finding a set of vehicles that broadcast their captured videos. The function for calculating priorities for each vehicle is given, and we find the set of vehicles with the maximum sum of priority values within the given bandwidth. The input of the problem is as follows.

- Set $Grid$ of $m \times m$ cells in intersection
- Set $Vehicle = \{v_1, \cdots, v_n\}$ of vehicles. Each vehicle v_i has the following attributes:
  - Speed [m/s]: $spd_i \geq 0$
  - Position: $pos_i = (x_i, y_i)$ x_i, y_i are real numbers
  - Cells captured by v_i Cap_i
 - Capturing direction dir_i: dir_i is measured by in-vehicle sensor.
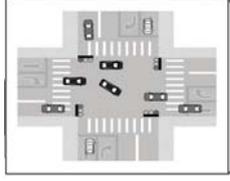
Figure 1. Bird's-Eye-View of Intersection
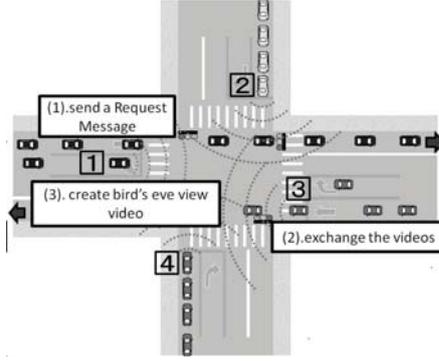
Figure 2. Overview of Method Procedures
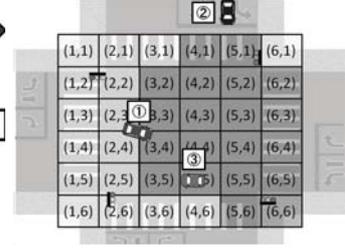
Figure 3. Cells in Intersection

– Video quality $qual_i$: real number between 0 and 1

- Video bit rate [Mbps] $BR_i \geq 0$
- Available bandwidth [Mbps] $W \geq 0$
- Priority function $Priority(v_i) \geq 0$
- The set of requesting cells in intersection $Req_i \subseteq Grid$

The output is the set $V_{send} \subseteq Vehicle$ of vehicles that broadcast their captured videos.

In order to compose a bird's-eye-view video of the intersection, some of the requested cells have to be captured by vehicles in V_{send}. Also, the sum of video bitrates must be within W. Additionally, the vehicles in V_{send} must have higher priority than other vehicles. Thus, the following constraints (1), (2) and (3) must hold.

$$Req_i \cap Cap_i \neq \emptyset \tag{1}$$

$$\sum_{v_i \in V_{send}} BR_i \leq W \tag{2}$$

$$\forall v_i \in V_{send}, \forall v_j \in Vehicle \backslash V_{send}: \\ Priority(v_j) \leq Priority(v_i) \tag{3}$$

We maximize the total priority of vehicles in V_{send}.

$$\mathbf{maximize} \sum_{v_i \in V_{send}} Priority(v_i) \tag{4}$$

3.4 Details of Proposed Method

In this section, we explain our method to solve the problem. The proposed method finds the set of vehicles that can capture the requested cells for each direction with good quality. The method consists of the following two phases.

- Exchange information on speed, direction, etc.
- Select vehicles that broadcast their videos

We explain the details of these phases below.

Information Exchange Phase
In this phase, vehicles exchange their information. Vehicles exchange *hello messages* regularly to share vehicle attributes. Since we are assuming that the speed of each vehicle is no more than 45[km/h] and the distance between moving vehicles is more than 5[m], one vehicle passes a position every 0.5[s] or more. Therefore, we set the broadcasting interval to 0.5[s]. A hello message contains the following items:

- $(i, spd_i, pos_i, Cap_i, dir_i, qual_i)$

Here i denotes a vehicle ID.

A vehicle that requests a video starts broadcasting request messages 100[m] before the center of the intersection. The request message contains the items of a hello message and the set of cells $Req_i \subseteq Grid$.

Vehicle Selection Phase
Each vehicle transits to this phase when it receives a request message. The set of vehicles each vehicle has received hello messages from is denoted $V_{share}(\subseteq Vehicle)$. The vehicle calculates the priorities of every vehicle in V_{share} from the request message and the hello messages, and retains them in a list. When the vehicle receives a new hello message, it updates the list of every vehicle's priority, selects a vehicle with the highest priority for each of four directions, and then adds the vehicle to V_{send}.

In our method, we make the following vehicles have high priority.

- Vehicles capturing larger number of requested cells. $f_{cap}(v_i)$ denotes the number of cells that Vehicle v_i captures and is calculated as follows.

$$f_{cap}(v_i) = \frac{(|Cap_i \cap Sub_j| + 2 \times |Cap_i \cap Req_j|)}{|Sub_j| + 2 \times |Req_j|} \times 10 \tag{5}$$

Here the set of eight-neighborhood-cells of a cell in Req_j is denoted as $Sub_j(\subseteq Grid)$.

- Vehicles closer to the center of an intersection. $f_{pos}(v_i)$ denotes the distance between v_i and the center of the intersection.
- Vehicles capturing higher quality video. $f_{qual}(v_i)$ denotes the quality of video that v_i captures. This value is normalized to the range $[0, 10]$.
- Vehicles with lower speed. $f_{spd}(v_i)$ denotes the speed of v_i.

We define the priority value assigned to a vehicle v_i as follows:

$$
\begin{aligned}
Priority(v_i) \quad = \quad & k_{cap} \times f_{cap}(v_i) + k_{pos} \times f_{pos}(v_i) \\
& + k_{qual} \times f_{qual}(v_i) + k_{spd} \times f_{spd}(v_i).
\end{aligned}
\tag{6}
$$

Here k_{cap}, k_{pos}, k_{qual}, and k_{spd} denote coefficient values for the functions respectively. According to the preliminary experiment about the coefficient values, we have confirmed that k_{pos} and k_{spd} have little effect on the result. Therefore, we decide that $k_{cap} = 1$, $k_{pos} = 0$, $k_{qual} = 2$, and $k_{spd} = 0$.

Each vehicle selects V_{send} by using the following algorithm to autonomously judge whether it is in V_{send} or not.

Step1. $V_{send} = \emptyset$

Step2. Calculate $Priority(v_i)$ for each vehicle v_i in V_{share}

Step3. Add the vehicle with the highest priority to V_{send} for each of four directions while the total bandwidth is less than W.

4 Experimental Results

In order to check the quality of the composed video and the bandwidth usage by the proposed method, we implemented our method on QualNet [2] simulator and conducted simulation-based evaluations.

4.1 Setting of Experiment

We created terrain data (142[m]×142[m]) to imitate an actual intersection in Kyoto, Japan. We set the traffic light for the direction of east and west to blue, and that for the direction of north and south to red. We reproduced a high vehicle density situation with 0 to 10[m] interval, and placed four vehicles in front of the red light. We made many vehicles run along the street extending east and west. As shown in Fig. 2, the simulation starts when vehicle 1, which is going to turn right, moved into the intersection. The simulation ends when there are no vehicles in the opposite lane and vehicle 1 is ready to turn right. Each vehicle was sending video at 600[Kbps] bit rate. We show the parameters of this experiment in Tables 1 and 2.

Table 1. Simulation parameters

the number of vehicles	60
Req	3 cells
Sub	12 cells
Wireless LAN standard	IEEE802.11b
updating interval of priority list	0.5 [s]

Table 2. Packet parameters

	packet size	sending interval
hello message	300 [byte]	0.5 [s]
request message	300 [byte]	0.5 [s]
video data packet	1500 [byte]	0.02 [s]

4.2 Evaluation Criteria

We show evaluation criteria in this study below.

(1) Packet arrival ratio of transmitted videos: In order to evaluate whether our method uses the communication bandwidth efficiently, we measure the packet arrival ratio of videos which vehicles in V_{send} sent to the request vehicle.

(2) Number of cells in of Req received videos cover: In order to evaluate whether the request vehicle can grasp its blind spots in the intersection, we measure the number of cells in Req which the received videos cover.

(3) Average quality of videos request vehicle received: In this experiment, we could not set actual quality of videos which vehicles captures because we evaluated it in a network simulator. So, in order to evaluate this criterion, we consider that the video which each vehicle captures is assigned to one of three grades (10 points, 5 points, and 0 point) at random.

(4) Average priority of the selected vehicles in V_{send}: In order to evaluate whether vehicles with high priority are selected in V_{send}, we measure the average priority of the selected vehicles in V_{send}.

4.3 Evaluation Method

We compared the following three methods of choosing the suitable vehicles that capture and send their video.

  (i) all vehicles in the intersection

 (ii) vehicles closest to the center for four directions

(iii) proposed method

4.4 Simulation Results

We show the simulation results of the proposed method and other two methods in Table 3.

Table 3. Comparison Results

	(i) All vehicles	(ii) Vehicles near to the center	(iii) Proposed method
(1)Packet arrival ratio[%]	9.9	88.4	89.0
(2)Num. of covered cells	0.59	1.57	1.63
(3)Avg. video quality	5.0	3.8	6.5
(4)Avg. priority in V_{send}	4.5	5.2	13.0

Evaluation Criterion (1): Packet arrival ratio of transmitted videos: As shown in Table 3, we see that the proposed method selected vehicles that use communication bandwidth effectively and send video data at a high probability with a fewer packet collisions than the other methods.

Comparing Method (ii) with the proposed method (Method (iii)), there were little difference of the packet arrival ratio. So, we see that the position of vehicles has little effect on this criterion when the density of vehicles is high like this experiment.

Evaluation Criterion (2): Number of cells in Req which the received videos cover Similarly to Criterion (1), we see that the proposed method selected vehicles which capture the requested cells more than the other methods. Also, Comparing Method (ii) with the proposed method (Method (iii)), there were little difference. This is because the density of vehicles is high in this experiment. As future work, we will evaluate the method when the density of vehicles is low enough.

Evaluation Criterion (3): Average quality of videos which the request vehicle received This result shows that the proposed method selected vehicles which have higher video quality than the other methods. On of the reasons is that the proposed method tries to preferentially select vehicles which can capture video in a high quality.

The result of Method (ii) is the worst among the three although the method is good as the proposed method in Criteria (1) and (2). As a reason of this, Method (ii) selects vehicles based on information about their position only.

Evaluation Criterion (4): Average priority of the selected vehicles in V_{send} We see that the proposed method selects vehicles with much higher priority than other methods. The reason is that, the proposed method selects vehicles taking into account not only their position but also cells which they capture and their video quality.

As a result, we confirmed that the proposed method selected vehicles that capture videos covering more cells requested by the request vehicle in higher quality.

5 Conclusion

In this paper, aiming to provide a system for preventing traffic accidents at intersections, we proposed an inter-vehicle communication protocol for composing videos captured from multiple vehicles at an intersection and sharing the composed video wirelessly among neighboring vehicles. Through simulations with a QualNet simulator, we confirmed that our method achieves good video quality for practical use.

As future work, we are planning to implement our method so that it works on real cars, and investigate effectiveness by using the implemented system in actual driving situations.

References

[1] ITS: "Ministry of Land, Infrastructure, Transport and Tourism,"

http://www.mlit.go.jp/road/ITS/.

[2] Scalable Network Technologies, Inc.: "QualNet,"

http://www.scalable-networks.com/.

[3] Honda Motor Co., Ltd.: "Honda Technology,"

http://www.honda.co.jp/news/2009/4090219.html?from=rss.

[4] Nissan Motor Co., Ltd.: "SAFETY ACTIVITIES,"

http://www.nissan-global.com/EN/SAFETY/.

[5] Ota, D., Shintaro, O., Ikeuchi, K.: "Visual Reconstruction of an Intersection by Integrating Cameras on Multiple Vehicles," *Conf. on Machine Vision Applications (MVA2007)*, pp.335–338(2007).

[6] Meng, G., Mostafa, H.A., Ellen, W.Z.: "V3: A vehicle-to-vehicle live video streaming architecture," *Proc. of the 3rd IEEE Int'l Conf. on Pervasive Computing and Communications (PerCom 2005)*, pp.171–180(2005).

[7] National Executive Committee: "SPACE-BASED POSITIONING NAVIGATION & TIMING,"

http://pnt.gov/.