Generating Bird's Eye View Images Depending on Vehicle Positions by View Interpolation

Jien Kato and Noritaka Sekiyama
Department of Systems and Social Informatics
Graduate School of Information Science
Nagoya University, Japan
jien@is.nagoya-u.ac.jp

Abstract

This paper proposed a method for generating bird's eye view images using multi-cameras set in high points at intersections. As the generated images follows own vehicle's motion and includes the vehicle itself, it enables the driver handing the surrounding situations conveniently. Experiments have shown the advantages of our method in save driving supporting.

1. Introduction

Recently in Japan, death rate of traffic accident had reduced every year. It shall be owed to the popularization of vehicle equipped safe-driving supporting system such as ABS, airbag and so on. However, the traffic accident was not decreased in quantity obviously. Every year, still a lot of people were injured by traffic accident. Presently, more attention is being paid on accident avoiding, one example is collision-prevent system. With more and more research be done in this aspect, it is becoming distinct that there is limitation for deal with such issue with single vehicle. In order to reduce the traffic accident in quantity, it is necessary to develop accident avoiding technology based on collaboration between vehicles and traffic infrastructures.

According to the traffic accident annual report [1], 72.6% of the accidents were caused by the driver acted unaware of front, inadvertence of motion or non-confirmation of safety. Furthermore, about 46.3% were occurred near intersection. That is because the complex environment of intersection lacks drivers' awareness.

This paper introduces a collaborative safe-driving supporting system. We assume plural cameras (called infra-camera) have been set in intersections. Vehicle receives image sequences that generated by each infra-camera in high view point while entering the intersections. The received multi-view images are then processed to form a bird's eye view image sequence that always follows the position/direction of the vehicle. We try to enhance the drivers' awareness by presenting such a real-time image sequence to them.

2. Approach

The main purpose of this research is providing information for safe-driving in complicated intersection environment. Therefore, we shall consider the following three issues.

1) As driver's perspective has blind spots, it is necessary to be enhanced.
2) For drivers realize and deal with dangerous based on the positional relationship between own vehicle and other vehicles or goers, it is necessary to provide clear information of around traffic situation to the drivers.
3) In order to support driver to judge instantly, the information shall be real-time.

To deal with issue 1, we think that providing images taken from a high point are efficient for the drivers. Since including some of the driver's blind spots, the perspective of a high point is broader than the driver's own perspective. Specially, in most case, the own vehicle's back or around and the other side of opposite vehicles are blind spots in driver's perspective, but there are visible in the high point perspective.

To deal with issue 2, in order to support driver to grasp the situation around, we take for third person's perspective as a solution. It is because the third person's perspective include own vehicle, this enable a visual representation of the positional relationship between own vehicle and other vehicles. Ordinarily, for the driver, scene through the front glass is changing depending on the own vehicle's position and direction, while the third person's perspective is independent. As a result, it is not easy for drivers to correspond third person's perspective with own vehicle. To deal with this, we provide our suggestive images from a virtual high point which always follows the vehicle's motion in order to make the driver feeling intuitive.

Our research is different from previous researches at the point that the virtual perspective keeps on following own vehicle's motion. As in research [2], a virtual mirror is proposed to reduce blind spots of images from high point camera. Whereas, the driver needs to correspond the suggestive images and own vehicle, and such correspondence is difficult to finish at a moment.

3. Vehicle--following bird's eye view image generating

In our method, we use real infra-camera images to generate virtual camera images as shown in Fig.1. In this way, seamlessly integrate the images from each camera, and generate corresponding bird's eye view video based on vehicle's position for the driver.

We suppose that the vehicle receive its position and direction information via vehicle-road communication when enter into the intersection environment. If the vehicle has been not yet captured from any camera, the suggestive images will be just the images output from the camera that has the same direction with the vehicle. In the other case, if the vehicle has been captured by cameras, the images from two cameras with the closest direction to the vehicle's will be selected to generate the images of the virtual camera by morphing. Finally, the images will be further processed into own-vehicle–centered images of bird's eye view, which follows own vehicle's motion.

In the following sections, we will explain the method of generating intermediate image in detail.

3.1 Vehicle's information

In our research, we assume that there are eight infra-cameras set around the intersection. We assign ID from 0 to 7 to them. The detailed arrangement is shown in Fig.1. The vehicle's position information is expressed in coordinates as *(xv,yv)* with centre of intersection as origin, and the direction is expressed as angle *θ(rad)* which between the vehicle's direction and the *ID=0* camera's direction. The infra-camera whose images are input to morphing as first image and final image are identified as ID_{src} and ID_{dst}, and could be calculated as follows:

$$IDsrc = \left[\frac{4\theta}{\pi}\right] \quad (1)$$

$$IDdst = \left[\frac{4\theta}{\pi}\right] + 1 \quad (2)$$

Here, [O] means the biggest integer that less than or equal to O, and $0 \leq \theta \leq 2\pi$.

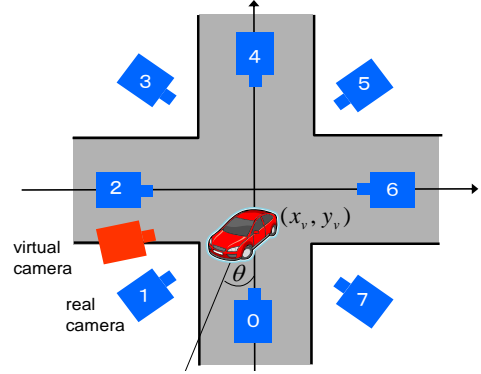### 3.2 Generate intermediate image by Morphing

Figure 1 Infra-cameras and the relation with the car.

3.2.1. Morphing rate. In order to generate the image of virtual perspective, we take morphing as our method. Morphing is a method of generating intermediate image by corresponding feature points of first image and final image. By setting different morphing rate, any intermediate image could be generated. We use the angle *θ(rad)* between the vehicle's direction and the ID=0 camera's direction to calculate the morphing rate *t* as follows.

$$t = \frac{4\theta}{\pi} - \left[\frac{4\theta}{\pi}\right] \quad (0 \leq t \leq 1) \quad (3)$$

3.2.2. Cross-dissolve. Cross-dissolve is a method of blending two pixels which have the same coordinates on the first and final image. With cross-dissolve, intermediate image could be generated by interpolating the colors using the colors of two input images. We use cross-dissolve to calculate the feature points' pixel value (RGB value) *c*.

$$c(t) = (1-t) \times c(0) + t \times c(1) \quad (4)$$

Here, *t* is morphing rate, $c(0)$ and $c(1)$ respectively correspond with pixel value from first and final image.

3.2.3. Warping. Warping is an operation to estimate mapping function for each pixel while transforming the first image to the final image. By using warping, we can generate the intermediate image that correctly gives the shape of objects in the intermediate viewpoint.

The feature point's position in intermediate image can be calculated by corresponding of its position in first and final images. The formula is written as

$$p(t) = (1-t) \times p(0) + t \times p(1) \quad (5)$$

Here, $p(t)$ is the coordinate of the feature point, $p(0)$ and $p(1)$ mean the coordinate in first and final image.

Then, we use the feature points of the intermediate image as a precondition to estimate the other points by corresponding of the first and final image. We use the relative coordinates of feature points to express these that are not feature points. To do this, we use a triangle
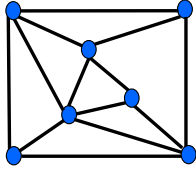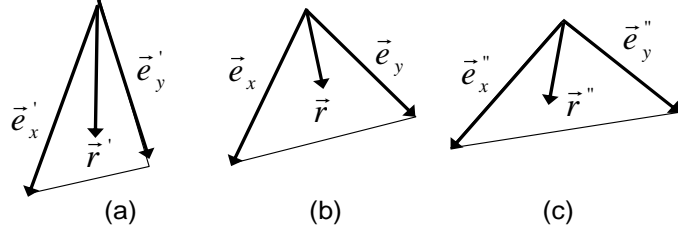
Figure 2 Mesh.

(a)　　　　(b)　　　　(c)

Figure 3 Relative coordinates.

- CPU: AMD Athlon(tm) 64 Processor 3000+ 2GHz
- RAM: 1GB
- Develop Environment: Microsoft Visual C++, OpenCV
- Network Camera: Panasonic BB-HCM100

Here, we simulate an actual situation which is always dangerous for drivers, and record the video of it. As shown in Fig.4. While own vehicle is about to make a right turn, a bike is approaching behind the opposite vehicle. In this situation, bike is in the blind spot of the driver's perspective. If the driver makes a right turn without recognition of the existence of the bike, the possibility of collision is very high. Here, we set the vehicles' speed as 10km/h-20km/h, the bike as running straightly in a speed of 40km/h and made the video. Due to the capability of communication with network cameras only can provide a frame rate of 2-5 fps, the finished video has a frame rate of 5fps.

4.2 Experiment results

We firstly separate each frame into different area of background, own vehicle, opposite vehicle and bike as well as manually input the corresponding feature points. And then, generate the output image according to our proposed method. At last, overlap them on the background to form the bird's eye view video. The resulting images are show in Fig.5 in time sequence.

4.3 Remarks

We firstly compare the suggestive video with the of driver's perspective. The driver's perspective which temporal corresponds to the frame 11 and frame 12 is shown in Fig.6. In the Fig.6, the left one is the opposite vehicle, and there is a bike approaching in the other side of it. The bike can't be seen until in frame 12. At that time (frame 12), there is not enough time left to the driver to make an operation to avoid the collision. In the other side, the bike was captured in the perspective of infra-camera much earlier. As shown in Fig.5, the suggestive video generated from infra-cameras can fix the blind spot of driver's perspective and provide an early reminder of the approaching bike. Accordingly, we think it is effective of reducing driver's blind spot.

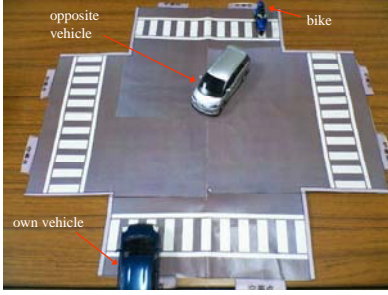Following, we consider the sensation that bought to

Figure 4 Situation with danger.

mesh with feature points as vertex which is shown in Fig. 2.

With one vertex as a reference, any point in the mesh can be represented by the base vectors corresponding with the two borders. We take $\vec{e}_x, \vec{e}_y$ for base vectors, $\vec{r}$ for any point's vector. The ratio of the base vector is identified as u and v. Then, we come with the following formula (Fig. 3(b)):

$$\vec{r} = u\vec{e}_x + v\vec{e}_y \quad (u,v \geq 0) \quad (6)$$

According to the formula shown above, we calculate u and v to each non-feature points.Then, estimate their relative coordinates on first and final image as shown Eq. (7) and (8) (see Fig. 3 (a) and (c)).

$$\vec{r}' = u\vec{e}_x' + v\vec{e}_y' \quad (7)$$

$$\vec{r}'' = u\vec{e}_x'' + v\vec{e}_y'' \quad (8)$$

Here, though $\vec{r}$, $\vec{e}_x$ and $\vec{e}_y$ are different in each image, the ration u and v are always the same. Therefore, put u and v that acquired from Eq. (6) into the Eq. (7) and (8), we can get the position of all the non-feature points from the first and final images.

4. Experiments

In this section, we will evaluate our method in three aspects: blind spots reducing, sensation for driver and real-time property.

4.1 Experiment Environment

As it is difficult to evaluate our method in actual intersection, we use an intersection model at 1:38 scale instead. Such a model simulates an intersection environment where each direction has one traffic lane, and road width is 0.95m. The experimental conditions are shown below:

drivers from the suggestive video. As the perspective of suggestive video is according to the vehicle's motion, it facilitates the driver to correspond the images with his own perspective. As in the suggestive video, vehicle becomes far away while it entered the intersection. To improve the sensation of following own vehicle, we can zoom in the focus to keep a changeless distance between virtual perspective and vehicle, but that will lead to reduction of the filed of vision.

The processing for generating bird's eye view image has not yet reached real time.

5. Conclusion

In this paper, we proposed a method for generating bird's eye view images using infra-camera images from a high viewpoint. As the generated images follows own vehicle's motion and includes the vehicle itself, it enables the driver handing the surrounding situations conveniently. Experimental results have shown the advantages of our method in save driving supporting. As the future work, we should reduce processing time and automatically extract feature points and further find their correspondence on different camera images.

Reference

[1] A.B. Smith, C.D. Jones, and E.F. Roberts, "Article Title", *Journal*, Publisher, Location, Date, pp. 1-10.
[2] Jones, C.D., A.B. Smith, and E.F. Roberts, *Book Title*, Publisher, Location, Date.

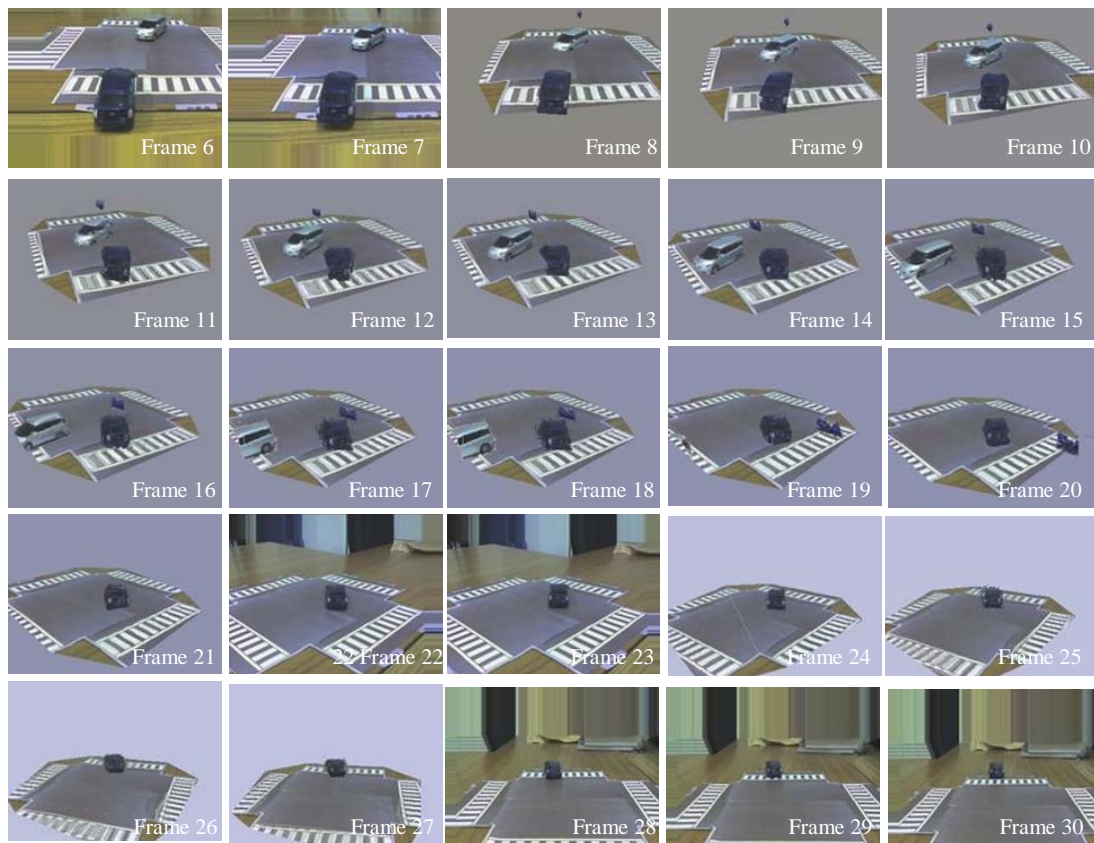Figure 6 Driver's perspective (left: frame 11, right: frame 12).

Figure 5 Generated bird's eye view which follows own vehicle's motion.