AR-Navi: An In-Vehicle Navigation System Using Video-Based Augmented Reality Technology

Yoshihisa Yamaguchi[1], Takashi Nakagawa[1], Kengo Akaho[2], Mitsushi Honda[2], Hirokazu Kato[2], and Shogo Nishida[2]

[1] Advanced Technology R&D Center, Mitsubishi Electric Corporation,
8-1-1, Tsukaguchi-honnmachi, Amagasaki, Hyogo, 661-8661, Japan
Yamaguchi.Yoshihisa@dp.MitsubishiElectric.co.jp,
Nakagawa.Takashi@dy.MitsubishiElectric.co.jp
[2] Department of Systems Innovation, Graduate School of Engineering Science,
Osaka University,
1-3, Machikaneyama, Toyonaka, Osaka, 560-8531, Japan
{akaho, honda}@nishilab.sys.es.osaka-u.ac.jp
{kato, nishida}@sys.es.osaka-u.ac.jp

Abstract. In this paper we propose AR-Navi, an in-vehicle navigation system that uses video-based augmented reality technology. AR-Navi includes an in-vehicle camera, which captures real-time video, and displays the guidance which overlays the video display. To solve several problems that arise when real-time video is used, we propose registration adjustment based on recognizing crosswalks and filtering images based on evaluated values. We have built and tested an AR-Navi prototype that uses the proposed methods and confirmed that the system can achieve registration accuracy for driver route guidance at intersections.

Keywords: Route guidance, In-vehicle navigation systems, Augmented reality.

1 Introduction

In recent years, over 24 million in-vehicle navigation systems have been produced in Japan [1] as such systems have rapidly become popular. The basic functions of these systems, such as map indication and route planning, are sufficiently powerful for current needs, so manufacturers are now seeking to improve in-vehicle navigation systems with regard to their user interface design. For example, new functions are being provided, such as displays showing a city landscape using 3D computer graphics (CG) or images combining aerial photos and existing maps.

Unfortunately, the usefulness of these functions is limited because of the following problems.

- 3D CG data and aerial photos are obtained in the past, so discrepancies might exist between the data and the present state.
- Existing data is created by sampling the real world to some limited degree of accuracy. This means data is simplified and thinned. For example, 3D CG data is

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part II, HCII 2007, LNCS 4558, pp. 1139–1147, 2007.

created according to the polygon number and images are created according to the resolution. Moreover, some data is deformed. That is, the data might not truly represent what a driver actually sees.

Because of such problems, a driver sometimes makes wrong turns or become confused, despite an in-vehicle system's correct route guidance, because the driver finds it difficult to make the correct correspondence between the real world and the displayed view.

In this paper, we propose AR-Navi, a system that uses video-based augmented reality technology. AR-Navi includes an in-vehicle camera, which captures real-time video, and displays the guidance which overlays the video display. Because the video is almost the same as what the driver sees, AR-Navi provides easily understandable guidance.

We give an overview of AR-Navi and the problems that limit existing systems in Section 2. We then propose solutions to these problems in Sections 3 and 4. We also describe an AR-Navi prototype and its performance.

Fig. 1. An example of route guidance using 3D CG

2 AR-Navi

AR-Navi is an in-vehicle navigation system that shows driver guidance overlaid on a video image captured from the driver's front view. Figure 2 shows a screen image of AR-Navi. A turn direction and turn location are shown by CG in this figure.

Other research has described in-vehicle navigation systems which use augmented reality technology. Sawada [2] describes the extraction of road geometry information and a way to draw CG roads on real-time video. Hu [3] describes the extraction of future points along a road and a way to adjust the location between the guidance drawn on the roads and real-time video. These studies have demonstrated the feasibility of high-precision alignment between real-time video and CG with image analysis. The image analysis technology used in these systems, however, has some limitations. First, the time complexity of the technology is high. Second, the technology is not robust regarding environmental conditions, such as rain, darkness, and backlighting. Thus, these systems are not used in current in-vehicle navigation systems. Our goal is to develop a commercial system which needs only a camera as additional equipment and is robust with respect to such environmental conditions.

Fig. 2. An AR-Navi image

2.1 Problems

We had to achieve the following to create AR-Navi in a way that overcame the problems affecting earlier systems:

- Suitable registration: A drivers feels uncomfortable when they encounter location gaps between the real-time video and the CG image.
- Information control: Too much guidance can lead to a driver not paying attention to traffic conditions.
- Elimination of inappropriate images: Captured images are sometimes unfit for overlapping guidance because the camera signal is blurred.

  In the following sections, we explain how we solved these problems.

2.1.1 Suitable Registration

When AR-Navi overlays information on a video image, it adjusts the positions of the information and the image based on the location and direction of the vehicle. If the location and direction contain errors, gaps will appear in the overlaid image.

Almost all current in-vehicle navigation systems calculate their own location and direction based on two steps [4]. First, the systems estimate their own location and direction from a combination of global positioning system (GPS) data and autonomous navigation data, such as that from a gyro and speed sensor. Second, they use a map-matching method which compares the result of the first step with road patterns. The systems correct the vehicle location if the driving record changes significantly (such as when the vehicle makes a turn) or there is a difference between the driving record and the road patterns. Immediately after a vehicle turn, the location and direction may show high accuracy but huge errors sometimes occur if the vehicle goes straight for a long time without a sufficiently strong GPS signal. We divide the gap between the location and the direction into three categories, and consider the resolution as follows.

- Back-and-forth gap: The longer a vehicle travels in a straight direction, the larger the back-and-forth gap becomes. Since a large back-and-forth gap can cause a driver to confuse a wrong intersection with the correct one, it is necessary to reduce this gap. (Problem 1)

- Right-and-left gap: Because map matching is used, the calculated location rarely goes off road. Thus, the right-and-left gap generally does not exceed the road width. If the vehicle does not leave the road, AR-Navi will not provide incorrect guidance so this gap is not a practical issue.
- Rotation gap: Current in-vehicle navigation systems estimate their own direction based on gyros. A rotation gap easily accumulates, though, because gyros integrate the change of the angle. Moreover, a measurement delay occurs that misaligns the guidance and video when the vehicle direction changes abruptly (such as when the vehicle makes a U-turn). It is therefore necessary to adjust the rotation gap. (Problem 2)

2.1.2 Information Control

Various safety guidelines state that in-vehicle navigation systems should provide required guidance quickly and effectively without interfering with driving [5]. For example, "The visual information to be displayed shall be sufficiently small in volume to enable the driver to comprehend it in a short time or shall be presented in portions for the driver to scan them in two or more steps."

If AR-Navi overlays CG guidance onto real-time video as a general AR application and provides the driver with overlaid guidance, the driver might need more time to recognize the guidance because he or she might be distracted by the interactive CG guidance and the background movie, which change in real time. Thus, we had to find a way to provide guidance without interfering with vehicle operation. (Problem 3)

2.1.3 Elimination of Inappropriate Images

Because of the camera settings and environmental conditions, captured images are sometimes unfit to be used for overlapped guidance due to problems such as

- Inappropriate brightness: Failure to control the lens aperture when, for example, moving from a shadowed area to a sunny area (Fig. 3(a)).
- Blur: Image blurring due to camera movement, especially right-and-left movement; for example, while turning at an intersection or while changing lanes (Fig. 3(b)).
- Inappropriate composition: The center of the target is located outside or on the edge of the screen (Fig. 3(c)).
- Limited visibility: The driver cannot see ahead because a large vehicle blocks his view (Fig. 3(d)).

If AR-Navi overlays guidance on images which correspond to any of the above conditions and provide driver with overlaid guidance, the driver will find it difficult to understand the real world information provided, so using real-time video has no merit in these situations. (Problem 4)

2.2 Proposed Methods

In this paper, we propose two techniques: (1) registration adjustment based on recognizing crosswalks, and (2) image filtering based on evaluated values. The first technique solves problem 1, and the second solves problems 2 to 4.

(a) Inappropriate brightness (b) Blur

(c) Inappropriate composition (d) Limited visibility

Fig. 3. Inappropriate images

3 Registration Adjustment Based on Recognizing Crosswalks

To decrease the back-and-front gap, we propose adjusting the current location by detecting intersections, which can be recognized using image analysis, that correspond with intersections on the map. The following conditions must be met to use such an algorithm in current in-vehicle navigation systems.

- Low time complexity: To allow its application in current in-vehicle navigation systems.
- High robustness: Robustness with regard to environmental conditions, such as weather and time, is necessary.

To satisfy the above conditions, this method finds crosswalks by applying image analysis to one line set on a captured image and located a certain distance ahead of the driver's vehicle. This line is called the scan line.

The back-and-front gap can then be calculated by comparing each intersection found with a corresponding intersection shown on a map, because it is assumed that an intersection lies between a pair of crosswalks within a limited distance.

This method has the following merits.

- It has low time complexity because the image analysis is aimed at a narrow area; i.e., the scan line.
- If the scan line is set near enough, other vehicles will rarely interfere with it while the vehicle is in motion, so the image analysis will be stable.

- If environmental conditions are bad, such as during rain and at night, the scan line is lit by the vehicle's own headlights.

This method is applied through the following steps:

1. Judge whether a crosswalk is present.
2. Store the position where a vehicle enters and leaves an intersection based on the crosswalks.
3. Judge whether a pair of crosswalks is appropriate.
4. If the crosswalk pair is appropriate, calculate the midway point between the location of entering an intersection and the location of leaving an intersection.
5. Adjust the back-and-front gap using the difference between the calculated midway point and the corresponding intersection on a map.

3.1 Judging Whether a Crosswalk Is Present

To find crosswalks, a brightness emphasis filter is applied to the brightness of each pixel in the scan line. Figure 4 shows the brightness emphasis filter. It is easy to get a highly accurate binarized result because the filter turns the raw brightness into the brightness shown in Fig. 5.

This method judges whether a crosswalk exists on the scan line by checking the number of white runs and black runs. If the length of each run is almost the same as the width of the white line of a crosswalk on the image and the white and black runs alternate, this method judges that an intersection has been found.

-1	0	0	0	0	2	0	0	0	0	-1

Fig. 4. Brightness emphasis filter

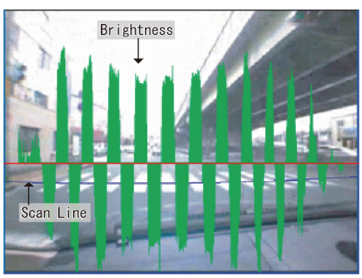

Fig. 5. Result of applying a crosswalk empathies filter

3.2 Judging Whether a Crosswalk Pair Is Appropriate

As explained in Section 3.1, to locate an intersection it is necessary to find two crosswalks. If two crosswalks are found, however, there is a possibility of image analysis failure. For this reason, this method does not adjust the location if two crosswalks do not constitute an appropriate pair; for example, if a crosswalk is too narrow or the distance between two crosswalks is too short.

4 Image Filtering Based on Evaluated Values

Image filtering is necessary to provide effective information and estimate whether images are inappropriate. Our approach to this is to clip still images from a real-time video stream. Thus, the driver is not distracted by a movie. However, if still images are clipped at only certain time intervals, the clipped images can be inappropriate. Moreover, even if clipped images are not inappropriate, there is a possibility that other clipped images would be more appropriate. Thus, in our system still images are clipped at all times, and the optimal image is selected from all clipped images according to certain evaluated values.

In this paper, image filtering consists of three steps:

1. Detection of inappropriate images
2. Selection of optimal images
3. Image updating

Thus, this process solves problems 2 to 4 stated above. In particular, problem 2 is reduced by step 1 because images including an excessive rotation gap are not selected. Problem 3 is solved through the clipping of still images and step 4. Problem 4 is solved by step 2, except under limited visibility.

4.1 Detection of Inappropriate Images

This step uses the following conditions to detect inappropriate images. The problem of limited visibility has not been solved in this prototype because the time complexity is increased due to the image analysis to locate vehicles. Other conditions are considered in this process.

4.2 Selection of Optimal Images

After the detection of inappropriate images, optimal images are selected based on evaluated values. We define an optimal image as one having the target intersection in the center, and this image is clipped when the distance between the vehicle and the intersection is reasonably short. Because two conditions are considered, the evaluated value is set as

$$E = k * p + (1 - k) * | D - D_{best} |$$

where k is constant ($0 <= k <= 1$), D is the current distance between the vehicle location and the target intersection, D_{best} is the ideal distance between the vehicle location and the intersection on the route guidance, and p is the coordinate distance between the intersection center and the image center.

The image having the lowest value of E is treated as the optimal image. In this research, we found that a suitable value of k is 0.4 and D_{best} is 30 m.

4.3 Image Updating

The updating frequency is important for safe driving. If the updating frequency is too short, the driver might be distracted by frequently changing images. Conversely, if the updating frequency is too long, the image will differ considerably from the real world. After careful consideration of the movement distance and elapsed time, we decided to have updating occur when the movement distance exceeded 25 m or the elapsed time exceeded 2.0 seconds.

5 Discussion

5.1 Prototype

We have implemented AR-Navi based on our in-vehicle navigation system, using a 320 * 240 pixel lens with a 6-mm focal length and 47° angle of view. When the user's vehicle is far from a target intersection, real-time video offers few benefits because the intersection is out of sight. Thus, the prototype shows the conventional full CG guidance when the vehicle is far from an intersection, and then shows guidance using the real-time video stream as the vehicle nears the intersection. Figure 6 is a screenshot from AR-Navi showing how AR-Navi provides guidance by using a triangular prism set on the intersection center and an indicator arrow.

Fig. 6. An AR-Navi screenshot

5.2 Discussion

We drove through an urban area to evaluate the guidance image provided by the prototype. AR-Navi provided turn guidance at 12 intersections, and the driver clearly recognized the target for all 12 intersections. The CG guidance was correctly overlaid on the intersection center on 9 of the 12 intersections. On the three other intersections, there was a short distance between the intersection center and the guidance, but these errors did not present a problem in actual use. We confirmed that this system achieved registration accuracy for turn-by-turn guidance by using the proposed methods: registration adjustment based on the recognition of crosswalks and image filtering based on the best-shot-selection process.

The target intersection was sometimes hidden by other vehicles, but since the system drew an indicator arrow in the part of the image corresponding to the sky (Fig. 6), these vehicles did not interfere with the overlaid guidance. We think the current display is practical for drivers in most situations, but we plan to further evaluate the display design and improve it.

During the test drive, the vehicle location error was sometimes increased by the registration adjustment. This happened when the system incorrectly matched an intersection of the map with one calculated by recognizing crosswalks. It will be necessary to solve this problem to improve the registration adjustment algorithm.

6 Conclusion

AR-Navi, a newly developed in-vehicle navigation system, shows driver guidance overlaid on a video image captured from the driver's front view. To solve several problems that arise when real-time video is used in such a system, we propose registration adjustment based on recognizing crosswalks and filtering images based on evaluated values. We have built and tested an AR-Navi prototype that uses the proposed methods and confirmed that the system can achieve registration accuracy for driver route guidance at intersections.

References

1. Japanese Ministry of Land, Infrastructure and Transport, Road Bureau ITS Homepage (2007) http://www.mlit.go.jp/road/ITS/
2. Sawano, H., Okada, M.: A Study on An Augmented Reality-based Rendering Method for Car Navigation Utilizing Real-time Video. In: FCV2004 - Tenth Korea-Japan Joint Workshop on Frontiers of Computer Vision, Fukuoka Japan (February 2004)
3. Zhencheng, H., Keiichi, U.: Solution of Camera Registration Problem Via 3D-2D Parameterized Model Matching for On-Road Navigation. International Journal of Image and Graphics 4(1), 3–20 (2004)
4. Quddus, M.A., Ochieng, W.Y., Zhao, L., Noland, R.B.: A General Map Matching Algorithm for Transport Telematics Applications. The GPS Solutions Journal 7(3), 157–167 (2003)
5. Japan Automobile Manufactures Association: Guideline for In-vehicle Display Systems Virsion 3.0