15th IAPR International Conference on Machine Vision Applications (MVA)
Nagoya University, Nagoya, Japan, May 8-12, 2017.

Image Recognition for Driver Assistance in Intelligent Vehicles

Hiroshi Murase
Nagoya University
Furo-cho, Chikusa-ku, Nagoya, 464-8603, Japan
murase@nagoya-u.jp

Abstract

Due to our aging society, driver assistance and automated driving have been intensely researched. Recognizing the surrounding environment utilizing image processing is a core technological element for driving intelligence. Specifically, technology must accurately detect pedestrians, obstacles such as other vehicles, condition of a driver, weather, etc. using cameras and various sensors. In this presentation, I introduce image recognition technology necessary for driver assistance.

1. Introduction

While driving, humans repeat actions that involve perception, decision-making, and operations. This set of actions is called driving behavior. To minimize the difference between the driver's behavior and the ideal behavior, driving intelligence calculates the ideal behavior based on the current surroundings and situations, and suggests information to the driver. In some cases, the system intervenes to achieve safer driving, reducing the operation load of the driver. For driving intelligence to determine the ideal driving behavior, it must recognize surrounding vehicles, pedestrians, road conditions, and blind spots. Then it must use this information to predict impending risks a few seconds in advance in order to calculate the trajectory of the vehicle's speed to reduce the risk. Numerous recognition technologies have been used to achieve this.

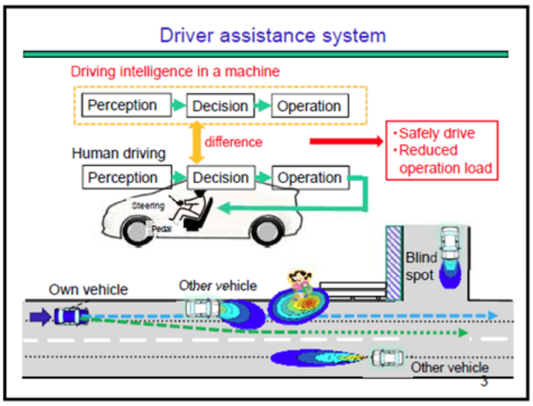

Fig.1 Driver assistance system

2. Driver assistance for the elderly

The proportion of car accidents caused by the elderly is increasing in Japan. More than half of such accidents are

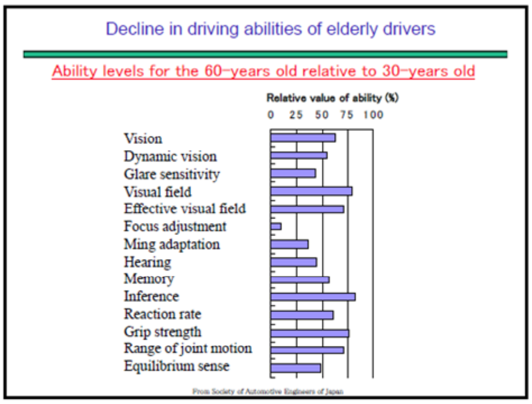

Fig.2 Decline in driving abilities of elderly drivers

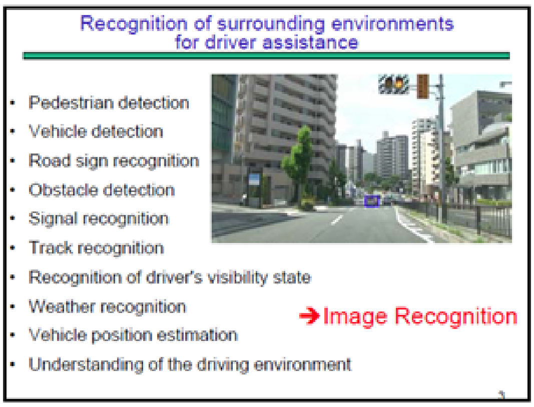

Fig.3 Recognition of surrounding environments for driver assistance

due to the delayed discovery of objects such as pedestrians, which is why assisting with recognition ability, especially image recognition, is important. Compared to younger drivers, elderly drivers have decreased levels of recognition, especially in the areas of focus adjustment, visual field range, and glare sensitivity. To overcome their decreased levels of abilities, recognition technology is becoming vital because it can discriminate objects in low-resolution imagery that is challenging for the most discerning human eyes. Image recognition technology necessary for driver assistance must be capable of pedestrian and vehicle detection, road sign recognition, obstacle detection, signal and track recognition, and identification of a driver's visibility state, weather conditions, and vehicle position.

3. Evolution of the recognition algorithm

Let's think about pedestrian detection. Within an input image, a system can focus on the local view from the

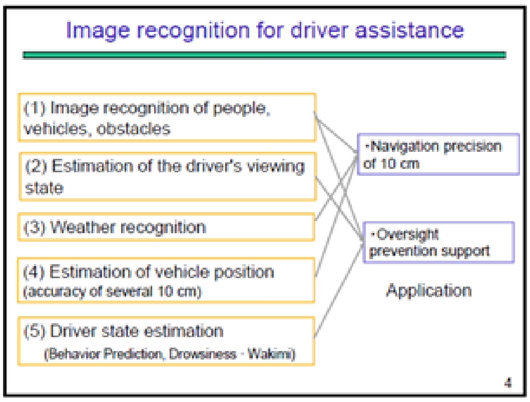

Fig.4 Image recognition for driver assistance

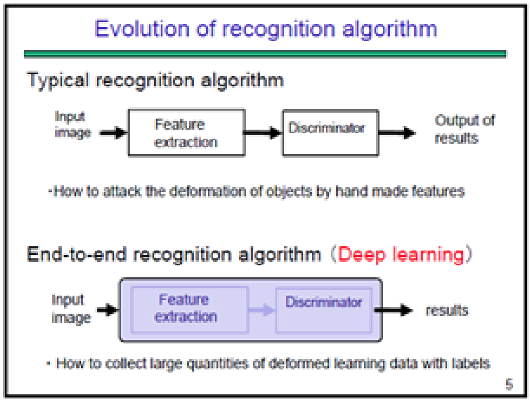

Fig.5 Evolution of recognition algorithm

window panel and discern pedestrians from the background. This can be achieved by sliding the window view into the full imagery system. Hence, the accuracy of the recognition algorithm to discern pedestrians from the background in the imagery input is crucial. This type of image recognition usually consists of two steps: feature extraction and a discriminator. Various methods and algorithms have been developed. For example for feature extraction, engineers have proposed employing features that are less affected by deformations and environmental light variations. The DPM [1] feature is one such method in which the system can learn the relative positions and movements of multiple parts of an object and match those to a pedestrian in the imagery. Thus, it can correctly recognize changing objects such as pedestrians. On the other hand, the end-to-end method, which learns the relationship between the imagery inputs and the output results with minimal consideration of the process, is receiving increased attention. This method does not distinguish between feature extraction and the discriminator. However, in recent years, the recognition functionality of deep learning in machine learning technology to find efficient methods to optimize the relationship between the inputs and the output results has drastically improved [2]. This method can also be viewed as machine learning for feature extraction. As feature descriptions and the quality of discriminator are becoming more sophisticated, the learning data are becoming more crucial. That is, collecting good quality data for learning is indispensable.

4. Collecting good learning data

In addition to recognition algorithms, the importance of the learning data is increasingly being acknowledged. To detect pedestrians, both positive samples showing pedestrians and negative samples with the backgrounds are necessary for learning. Good learning data should include diverse deformities and variations similar to real environments as well as correct labels. Gathering numerous learning samples can realize the former condition, whereas the latter requires that a correct label be attached to each sample. Because this results in contrasting conditions, various tricks need to be invented.

Methods to collect vast amounts of data can either artificially generate image data or automatically gather data from real environments. There are several methods to generate image data. Data Augmentation simply adds deformations to existing imagery to increase the number of data. Model-based Generation creates realistic data by applying a statistical model of image variation factors. Computer graphics program is used to generate image data.

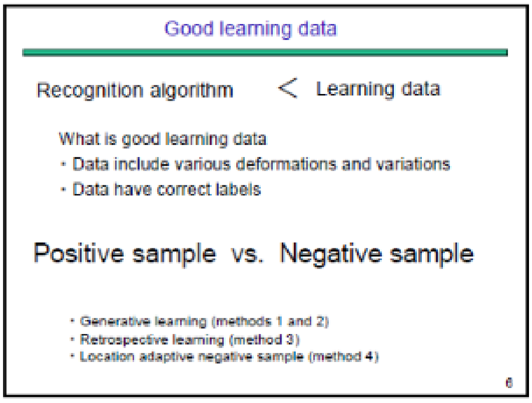

Fig.6 Good learning data

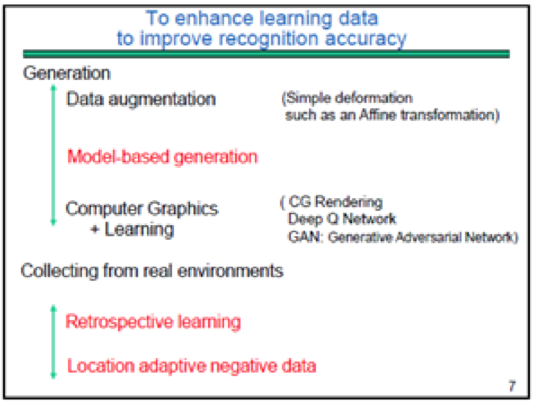

Fig.7 To enhance learning data to improve recognition accuracy

On the other hand, a key technology to collect image data from real environments is to incorporate a method to automatically add labels to some of the images. In this presentation, I introduce four methods that we tried: 1) Application of the degeneration factor to model real data samples and produce new data. 2) Application of augmentation or statistical variations to several learning data to generate new data. 3) Collection of time sequential learning data from real environments. 4) Collection of various negative samples from the locations on a map.

4.1 Generating learning data using models

As an example, consider a street sign. The image of a street sign captured by an in-vehicle camera may be deformed or deteriorated due to low resolution, optical blur, or motion blur caused by a slower shutter speed relative to the vehicle movement. Each negative factor can be modeled individually, and various parameters at a particular time can be statistically calculated using a genetic algorithm. The proposed method generates deteriorated and deformed images by applying the model to high-quality imagery. If diverse street sign images generated by the models resemble those in real environments, poor-quality images can be used as learning samples to improve the recognition accuracy of the deteriorated and deformed street signs [3].

4.2 Augmentation and deformation

In the case of street signs, the range of deformations is limited by the object's rigidity. However, the range of changes is much wider and the types of changes are more diverse for pedestrians because their shapes and the texture of clothing greatly vary. Accordingly, the proposed method deforms pedestrian image samples based on a statistical shape model of pedestrians and adding random clothing textures to generate new data after collecting a certain number of image samples. To be more specific, we combine two methods to generate a variety of pedestrian image data; one applies a statistical shape model created by extracting the outer forms of pedestrians and morphing the shapes, while the other applies Delaunay triangulation

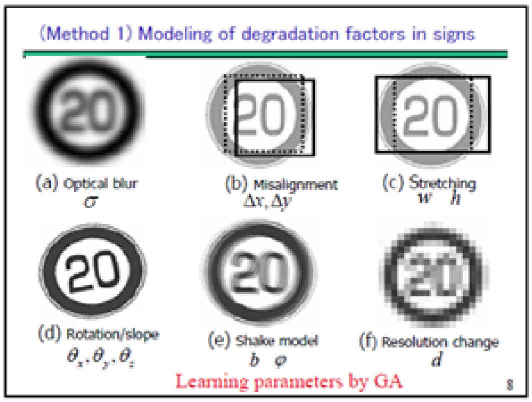

Fig.8 Modeling of degradation factors in sign

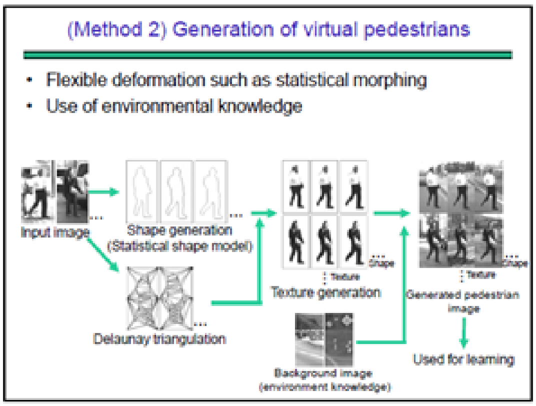

Fig.9 (Method 2) Generation of virtual pedestrians

to the human body texture and form. These methods can improve recognition accuracy [4].

4.3 On-line learning sample collection

Another method is to automatically collect learning samples from real imagery. Because street sign detection is difficult in deteriorated distant images, automatically collecting such samples is difficult. However, detecting street signs in close-up images is easy, and tracking already identified street signs is simple. Therefore, street sign images can be detected in close-up pictures and then working backwards in time, lower-quality distant images can be collected. By adding such images as learning samples, the recognition accuracy can be improved [5].

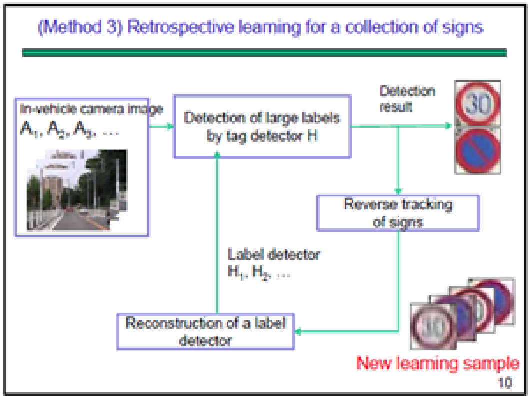

Fig.10 (Method 3) Retrospective learning for a collection of signs

4.4 Collecting negative samples at various locations

The recognition accuracy can be determined by the relationship between positive learning samples and negative learning samples. Negative learning samples are background imagery. The identifiable characteristics of the background differ greatly by location. However, GPS can provide vehicle location information. Thus, location adaptive image recognition can be established by collecting negative learning samples at various GPS-provided locations, improving the recognition accuracy more than employing a single location recognition method [6].

5. Detecting general obstacles

Obstacle recognition based on machine learning cannot detect objects not learned previously. In other words, it cannot recognize unexpected objects. Driver assistance requires recognition of hard-to-learn obstacles, for example, a ball on the road, a fallen tree limb, or cardboard boxes, whereas driving intelligence requires recognition of general obstacles, including hard-to-learn obstacles. The proposed method [7] stores previously collected image data from an in-vehicle camera in a database. This database is used to determine the difference between the current image and the stored images to detect an obstacle. Because a vehicle typically drives on a limited number of roads, creating an image database of often-travelled roads

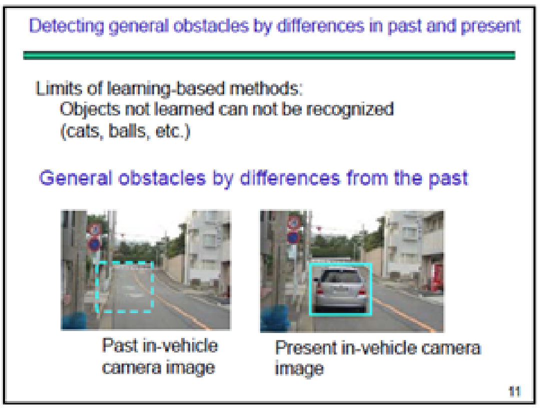

Fig.11 Detecting general obstacles by differences in past and present

is not difficult. However, the precise trajectory and driving speeds differ each time. Consequently, the location of the images must be matched to calculate the difference. Moreover, the lighting condition can vary as the car travels at different times of the day. Thus, an algorithm for finding differences must be robust against variations in brightness. Next, I discuss some examples detecting vehicles and traffic cones on the road.

6. Using image processing to detect objects that are easily overlooked by a driver

Advances in image recognition technology allow an in-vehicle camera to detect many objects such as pedestrians and other vehicles in the surrounding environment. Nevertheless providing every recognition result will disturb the driving, and desensitize the driver to risk. Accordingly, alerting the driver to objects that are most likely to be overlooked or dangerous after automatically sorting based on risk is desirable. Let's use the term visibility to indicate the likelihood of an object being overlooked. Here, I outline technologies that can estimate the visibility using image processing. Humans tend to overlook an object if another object is more prominent in the same view, while the peripheral region of the visual field of humans is more sensitive to movements than the shapes of the objects. Considering the characteristics of human visual field, a method to recognize the objects is proposed using regression [8, 9].

7. Recognizing weather

Changes in weather greatly affect a driver's visibility. For example, visibility decreases when the windshield is covered with raindrops or it is foggy outside. Thus, it is necessary to detect the condition of the windshield or the fog density. If a system can identify rain drops on a windshield, it may automatically start the windshield wipers. Rain drops can be identified by extracting their unique properties [10, 11]. In addition, if the fog density can be detected, fog lights can be turned on automatically as the fog density increases. The fog density can be

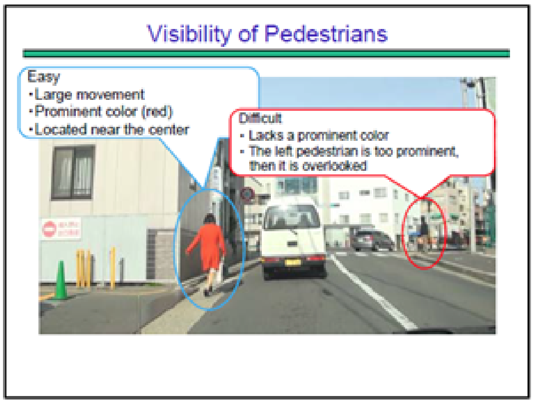

Fig.12 Visibility of pedestrian

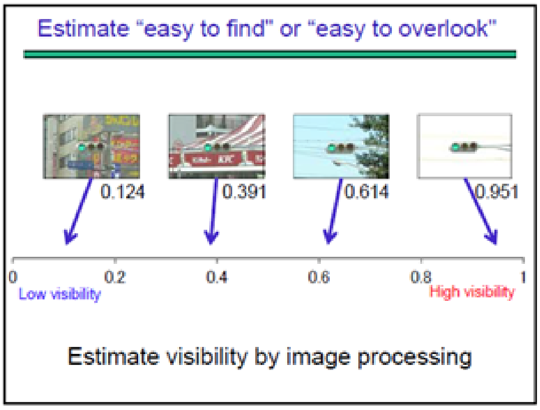

Fig.13 Estimate "easy to find" or "easy to overlook"

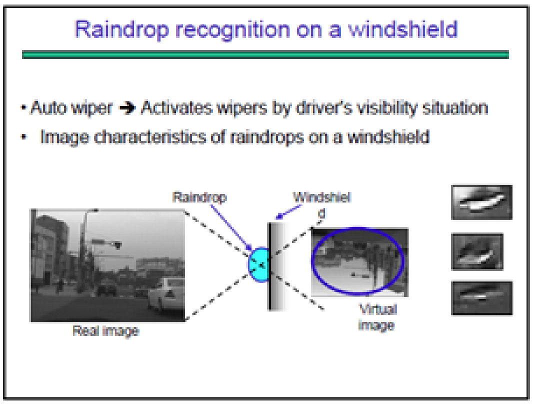

Fig.14 Raindrop recognition on a windshield

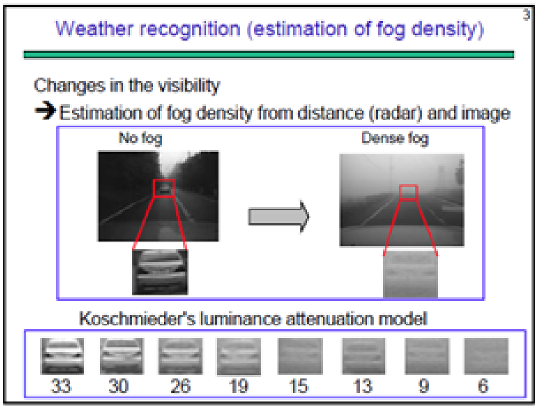

Fig.15 Weather recognition (estimation of fog density)

determined by extracting the scattering coefficient parameters from image processing based on Koschmieder's luminance attenuation model [12].

8. Estimating vehicle position

To guide a vehicle into the most fitting trajectory, the vehicle location must be accurately estimated on a map, especially when dealing with a vehicle in a lane where the location accuracy must be within 10 centimeters. GPS is a convenient location estimator, albeit with errors. In addition to GPS, methods using a camera or LIDAR are being researched. For instance, 3D LIDAR data and previously photographed map data can be matched to accurately estimate the vehicle location. SLAM technology, which simultaneously estimates the vehicle location and constructs a map, is becoming more prevalent. In general, SLAM uses the same sensor to create a map and estimate the vehicle location [13]. Here, I also introduce methods that use map data obtained from different sensors to match and estimate the vehicle location accurately. The first method uses an aerial image as the map image and to match the image of the forward view to estimate its location. Due to difference in the camera angles, geometric transformations and extraction of local characteristics are used to estimate the vehicle location. An experiment showed that the accuracy is within 20 cm, but aerial images are too expensive to obtain frequently. Thus, we devised another method using the in-vehicle camera to complement the aerial image [14].

The second example estimates the vehicle location by matching map data obtained by an omni-directional camera such as Google Street View to the in-vehicle camera image of forward view. In this example, the types of cameras differ, which may cause an issue where the images only partially match. To resolve this, I introduce the DP matching method as well as research on image quality improvement of map database by repeatedly generating images using omni-directional cameras [15, 16].

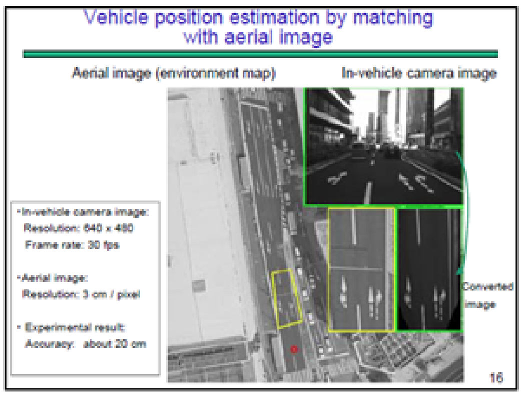

Fig.16 Vehicle position estimation by matching with aerial image

9. Predicting a driver's behavior in advance

If the system can predict a driver's behavior in ad-

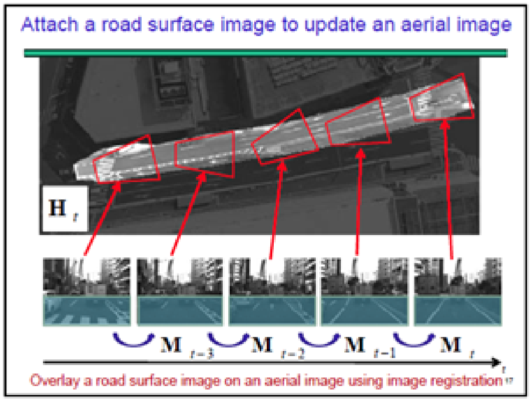

Fig.17 Attach a road surface image to update an aerial image

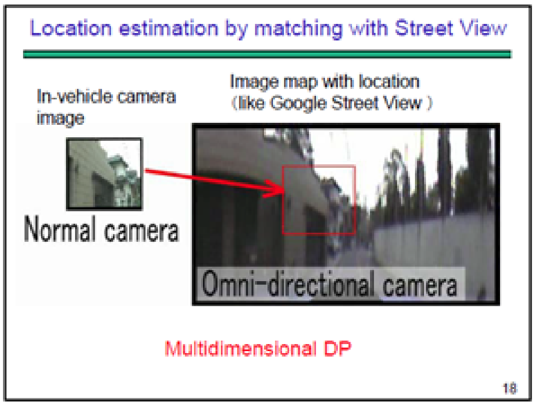

Fig.18 Location estimation by matching with Street View

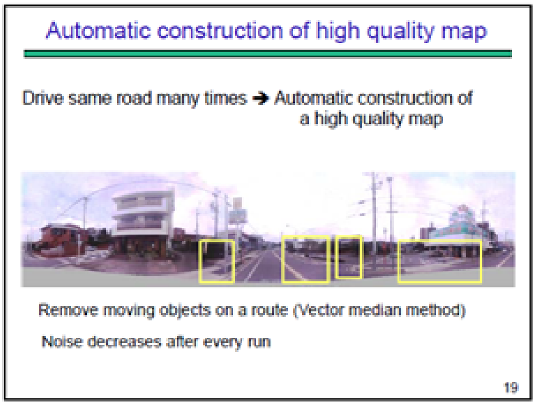

Fig.19 Automatic construction of high quality map

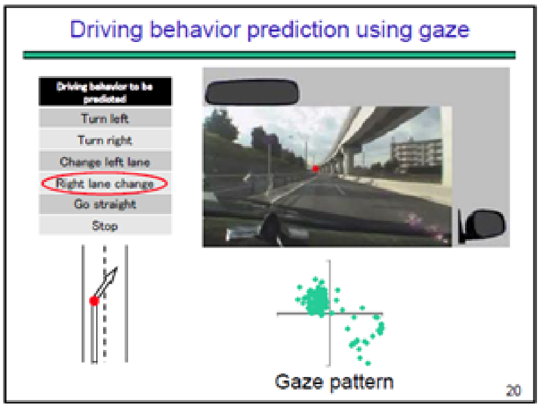

Fig.20 Automatic construction of high quality map

vance, a more suitable driver assistance is possible. Information on a driver’s handling operations or brake usage may only give the intended driving behavior immediately before the action occurs. However, a driver’s gaze information may predict the intended driving behavior a few seconds in advance. For example, a driver checks the other lane a few seconds before changing lanes. In this section, I’ll introduce a method to predict a driver’s behavior using both eye movements and vehicle data [17].

10. Conclusion

Recognition of the surrounding environment is a core element necessary to realize driving intelligence. In this presentation, I introduced a number of methods to recognize the surrounding environment. In the future, driver assistance will become even more sophisticated, and both high accuracy image recognition technologies and new types of recognition technologies that we have yet to imagine will be necessary.

References

[1] Pedro Felzenszwalb, David McAllester, Deva Ramanan, “A Discriminatively Trained, Multiscale, Deformable Part Model,” Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR2008), 2008/06/23

[2] Hiroshi Fukui, Takayoshi Yamashita, Yuji Yamauchi, Hironobu Fujiyoshi, Hiroshi Murase, "Pedestrian Detection Based on Deep Convolutional Neural Network with Ensemble Inference Network," IEEE Intelligent Vehicles Symposium (IV), pp.223-228, 2015/06/29

[3] Hiroyuki Ishida, Tomokazu Takahashi, Ichiro Ide, Yoshito Mekada, Hiroshi Murase, "Identification of Degraded Traffic Sign Symbols by a Generative Learning Method," Proceedings of the 18th IAPR International Conference on Pattern Recognition (ICPR2006), Vol.1, pp.531-534, 2006/08/01

[4] Hidefumi Yoshida, Daichi Suzuo, Daisuke Deguchi, Ichiro Ide, Hiroshi Murase, Takashi Machida, Yoshiko Kojima, "Pedestrian Detection by Scene Dependent Classifiers with Generative Learning," Proceedings of 2013 IEEE Intelligent Vehicles Symposium (IV2013), pp.654-659, 2013/06/25

[5] Daisuke Deguchi, Keisuke Doman, Ichiro Ide, Hiroshi Murase, "Improvement of a traffic sign detector by retrospective gathering of training samples from in-vehicle camera image sequences," Proceedings of Computer Vision in Vehicle Technology: From Earth to Mars (CVVT2010), 2010/11

[6] Daichi Suzuo, Daisuke Deguchi, Ichiro Ide, Hiroshi Murase, Hiroyuki Ishida, Yoshiko Kojima, "Environment adaptive pedestrian detection using in-vehicle camera and GPS," Proceedings of the International Conference on Computer Vision Theory and Applications (VISAPP) 2014, pp.354-361, 2014/01/07

[7] Haruya Kyutoku, Daisuke Deguchi, Tomokazu Takahashi, Yoshito Mekada, Ichiro Ide, Hiroshi Murase, "On-road Obstacle Detection by Comparing Present and Past In-vehicle Camera Images," Proceedings of IAPR Conference on Machine Vision Applications (MVA) 2011, pp.357-360, 2011/06/14

[8] Masashi Wakayama, Daisuke Deguchi, Keisuke Doman, Ichiro Ide, Hiroshi Murase, Yukimasa Tamatsu, "Estimation of the human performance for pedestrian detectability based on visual search and motion features," Proceedings of the 21st IAPR International Conference on Pattern Recognition (ICPR2012), pp.1940-1943, 2012/11/14

[9] Ryunosuke Tanishige, Daisuke Deguchi, Keisuke Doman, Yoshito Mekada, Ichiro Ide, Hiroshi Murase, "Prediction of Driver's Pedestrian Detectability by Image Processing Adaptive to Visual Fields of View," Proceedings of 2014 IEEE Conference on Intelligent Transportation Systems (ITSC2014), pp.1388-1393, 2014/10/10

[10] Tomokazu Takahashi, Kenji Mori, Hiroyuki Kurihata, Ichiro Ide, Hiroshi Murase, Takayuki Miyahara, Yukimasa Tamatsu, "Rain and Fog Recognition System Using Multiple In-vehicle Sensors," Proceedings of the 14th World Congress on Intelligent Transport Systems, 2007/10

[11] Hiroyuki Kurihata, Tomokazu Takahashi, Yoshito Mekada, Ichiro Ide, Hiroshi Murase, Yukimasa Tamatsu, Takayuki Miyahara, "Raindrop Detection from In-Vehicle Video Camera Images for Rainfall Judgment," Proceedings of the 1st International Conference on Innovative Computing, Information and Control (ICICIC2006), Vol.2, 2006/08

[12] Kenji Mori, Terutoshi Kato, Tomokazu Takahashi, Ichiro Ide, Hiroshi Murase, Takayuki Miyahara, Yukimasa Tamatsu, "Visibility Estimation in Foggy Conditions by In-vehicle Camera and Radar," Proceedings of the 1st International Conference on Innovative Computing, Information and Control (ICICIC2006), Vol.2, 2006/08

[13] David Robert Wong, Daisuke Deguchi, Ichiro Ide, Hiroshi Murase, "Single Camera Vehicle Localization Using Feature Scale Tracklets," IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, E100-A, No.2, pp.702-713, 2017/02/01

[14] Masafumi Noda, Tomokazu Takahashi, Daisuke Deguchi, Ichiro Ide, Hiroshi Murase, Yoshiko Kojima, Takashi Naito, "Vehicle Ego-localization by Matching In-vehicle Camera Images to an Aerial Image," Proceedings of Computer Vision in Vehicle Technology: From Earth to Mars (CVVT2010), 2010/11

[15] Hiroyuki Uchiyama, Daisuke Deguchi, Tomokazu Takahashi, Ichiro Ide, Hiroshi Murase, "Ego-localization using Streetscape Image Sequences from In-vehicle Cameras," Proceedings of 2009 IEEE Intelligent Vehicles Symposium (IV2009), pp.185-190, 2009/06

[16] Hiroyuki Uchiyama, Daisuke Deguchi, Tomokazu Takahashi, Ichiro Ide, Hiroshi Murase, "Removal of Moving Objects from a Street-View Image by Fusing Multiple Image Sequences," Proceedings of the 20th IAPR International Conference on Pattern Recognition (ICPR2010), pp.3456-3459, 2010/08/01

[17] Tatsuki Kamisaka, Masafumi Noda, Yoshito Mekada, Daisuke Deguchi, Ichiro Ide, Hiroshi Murase, “Prediction of driving behavior using driver's gaze information”, IEICE, PRMU, 111(47), pp.105-110, 2011/05/20