Self-Position Estimation Technology using Multi Angle Vision™

Wataru HASEGAWA Toshihiro MATSUMOTO Yuichi SUGIYAMA

Masayuki KISHIDA

Abstract

The needs for parking are high among the development for autonomous driving. DENSO TEN has proceeded with system development related to automatic parking using Multi Angle Vision™. An assistance system which partially supports parking operation already appears in the market now. In future, as automation moreover advances, the automation will expand and evolve from the parking system including autonomous driving (home parking, autonomous valet parking and others) into the automatic parking which can apply to the parking places even with ordinary vehicles. In accordance with this evolution, high-accuracy self-position estimation technology as well as more precise vehicle control is required simultaneously.

We assume application of Simultaneous Localization and Mapping (SLAM), the technology which estimates the self-position with the image measurement technology and creates the map simultaneously, to Multi Angle Vision™, and we propose the measure for the problem of positional accuracy affected by the synchronizing gap between cameras.

1. Introduction

The needs for parking are high among the development for autonomous driving. DENSO TEN has proceeded with system development related to automatic parking using Multi Angle Vision™. It is predicted that the current parking-assistance system will be expanded into a system for automatically parking a vehicle such as fully automatic parking at home and Autonomous Valet Parking, and home parking in future.

The future automatic parking needs not only a parking storage, but also an automatic driving to the storage place. Considering the situations of general parking space, high-accuracy vehicle control is needed because the high-density storage by Autonomous Valet Parking and home parking in a narrow site should also be taken into consideration. Toward the implementation, high-accuracy self-position estimation is required and LiDAR (Light Detection and Ranging) is primarily used to measure a position in a current automatic driving experiment. Since this is expensive despite the high measurement accuracy and resolution, it takes time to be widely adopted. On the other hand, Multi Angle Vision™ of DENSO TEN product is acknowledged as a monitoring system capable of monitoring all-surroundings of a vehicle without blind angle, and the number of vehicles mounting it has been gradually increasing. Utilization of this as a sensor for positional measurement has a greater advantage also for cost merit.

The image measurement technology and map creation technology are mostly used to measure self-position by image. The technology, which performs self-position estimation and a map creation of the surrounding area from images simultaneously, has been mounted on a robot and toy, and also utilized into various applications such as AR/VR technology and 3D object measurement. On the other hand, the front monitoring and object recognition have been mainly the mainstream for automotive applications, and the utilization of self-position estimation for parking applications is the technology from now on.

There are also utilization issues with image measurement technology and map creation technology by the camera of Multi Angle Vision™. Since the shutter timing of the camera is not synchronized especially for Multi Angle Vision™, it is predicted that positional error will be caused. Solutions to these issues are shown and self-position estimation technology by the future Multi Angle Vision™ is described in this article.

1.1 Future Automatic Parking System

Currently, various functions which support parking have started to be mounted on vehicle. Most of them support functions which steers a vehicle to a parking lot position and operates a steering wheel instead of a driver with grasping a parking frame and peripheral state.

It is predicted that parking-assistance system will gradually evolve to deal with complicated process starting from this point. **(Fig. 1)**

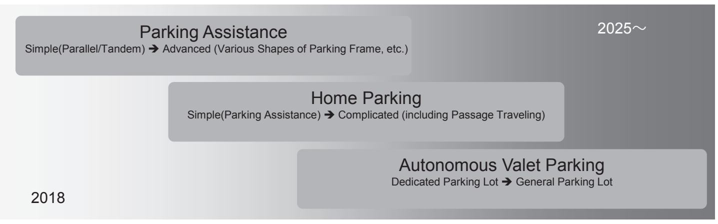

Fig. 1　Parking Functional Evolution

The next step is considered to be home parking. The entry and exit of a parking lot are essential operations in daily vehicle operation. However, it may be very difficult to park a vehicle in urban areas and dense residential areas due to narrow housing area and the like. For example, there are various difficulties such as a narrow road, heavy traffic, and obstacles (telephone pole and sign, etc.) near the entrance and the like. For this reason, home parking may cause great stress for drivers.

Home parking system is a system for automating storage, which also enables home parking with a high degree of difficulty to be stored smoothly and thus leads to the load reduction of the driver. **Fig. 2** shows an image of home parking. The driver starts the parking system after once stopping at a place where it is easy to park. The system performs parking operation while measuring self-position, calculating the trajectory to the storage place, and checking the surrounding state.

In urban areas and densely populated residential areas, the distance from a general road to the storage place may be shorter as in the **Fig. 2**. However, autonomous driving is simultaneously used to be applicable to the case with a relatively long distance. For home parking, a sufficient space may not be secured in a passage and parking area. Therefore, it is important to accurately grasp own vehicle position.

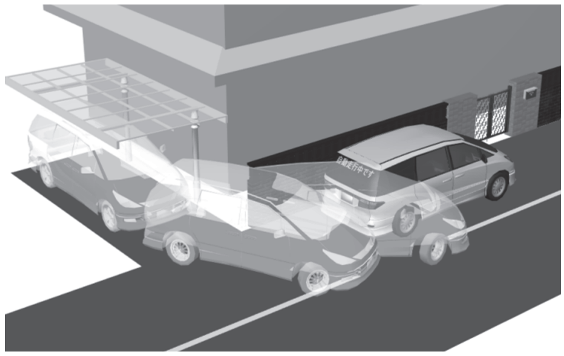

Fig. 2　Image of Home Parking

We assume that the next step is Autonomous Valet Parking. Valet parking much seen at overseas hotels is a service that a servant or valet receives customers' vehicles at a getting-off position, stores them, and moves them again to a getting-on position such as an entrance when their vehicles leave. Autonomous Valet Parking system is the one that is automated by autonomous driving technology and automatic parking technology. All the following movements, and storing operation to a parking area, and moving to a getting-on position in the wake of receiving a call are operated under the system control after the control of a vehicle is borrowed from a user at a getting-off position. Currently, research has been conducted by vehicle manufactures and national project.

Fig. 3 shows an image of Autonomous Valet Parking system. After all passengers get off the center vehicle, the control is given to the system to let the vehicle run and park automatically.

The operation which stores only applicable vehicles at a dedicated parking lot is launched in an early stage, and then we assume that the operation at a general parking lot where there are non-applicable vehicle and a pedestrian will be launched eventually.

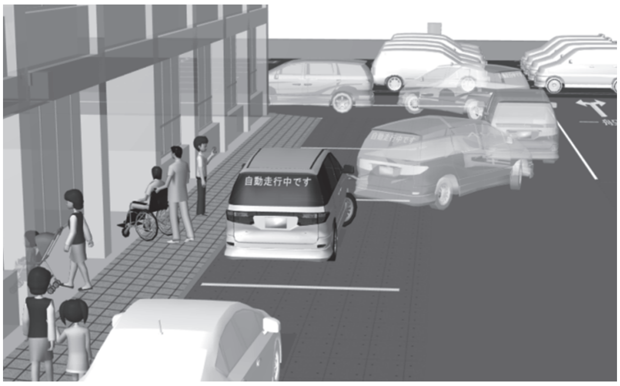

Fig. 3 Image of Autonomous Valet Parking

Storage in the dedicated parking lot, any user doesn't get on/off a user's vehicle in the storage, where any getting on/off space for vehicle surroundings isn't required to be secured, therefore it is being expected to store vehicles in higher density than a general parking. In order to realize this, more accurate vehicle positional control is an important issue.

In this way, sensing the surrounding environment and grasping its own vehicle position accurately can be considered as the important issues in the development of the future parking system.

2. Self-Position Estimation by Multi Angle Vision™

We think it is appropriate to apply the image measurement technology for self-position estimation in Multi Angle Vision™ with camera. Multi Angle Vision™ has four cameras to monitor the surroundings of vehicle and is arranged to reduce blind spots to a minimum. The improvement of self-position estimation accuracy even under the complicated environment such as home parking as well as the road capable of easily grasping the surrounding environment can be expected by performing a self-position estimation and a map creation in four directions simultaneously. In addition, it can also be expected that plural directions of the camera's optical axes may lead us to the improvement of the robustness against optical environment even when the surrounding state can't be grasped due to halation and the like.

2.1 Measurement Principle of the Image Measurement Technology

Stereo method and motion stereo method are overviewed as an example of the image measurement technology.

Three-dimensional feeling obtained by showing parallax images each independently to the left eye and the right eye enabled Wheastone (1802-1875) to invest the stereoscope. (This later became the origin of the word “stereo” for acoustic, etc.) Triangulation is the one that makes this principle developed to measure the distance direction to an object.

For a triangle on a plane, the vertex position can be obtained by determining the both angles of base ends and the base (stereo method).

This means that the direction of the same object is measured by two cameras (stereo camera) placed in parallel to have a view of the same area. Since the distance between cameras is known in advance, the distance and direction to the object are obtained. **(Fig. 4)**

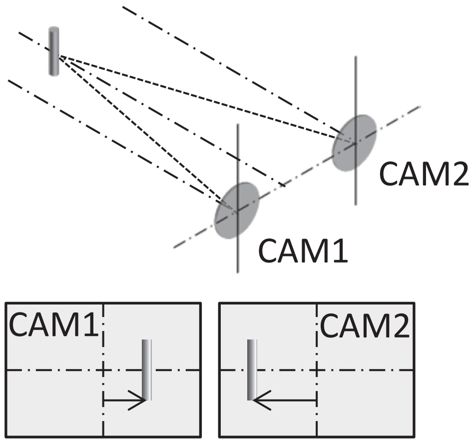

Fig. 4 Stereo Method

According to the same principle, the position is changed and photographed twice by one camera, which enables the coordinate to be obtained in the same way. This is called motion stereo method because one camera moves and applies the stereo method. **(Fig. 5)**

That is known in triangulation of the stereo method because the angles of the both ends are measured and the length of the base is fixed by a camera. On the other hand, the motion stereo method requires a movement, and the movement amount is the length of the base. However, any value is normally used at the initial stage because the moving distance and direction is undefined. Therefore, the coordinate value is the relative value. In order to convert this into an absolute value, the scale factor must be obtained by matching to the known map.

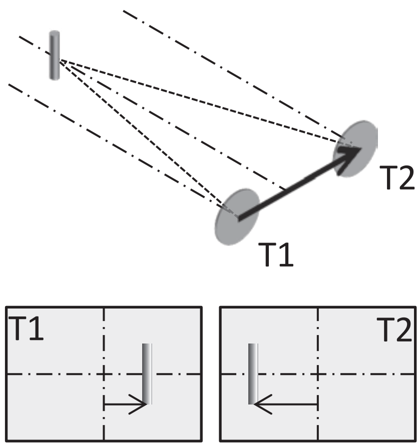

Fig. 5　Motion Stereo Method

Important issue in the stereo method and the motion stereo method is to choose the same target from two images. Thereby, some feature places are searched from the images and the matching process of the both is performed. This place is called a feature point and the value related to the match is called feature amount.

Based on the combined feature point, it is converted into a three-dimensional coordinate by the stereo method. A large number of three-dimensional coordinate points can be obtained from one frame image. Accumulating and saving those targets' coordinates enable a map of surrounding area to be created (map creation technology). Self-position can be localized by calculating backward from the map simultaneously. This is generally called Simultaneous Localization and Mapping (SLAM). The basic structure of SLAM using the motion stereo method is shown. **(Fig. 6)** The previous feature point is quoted to perform the comparison of the movement.

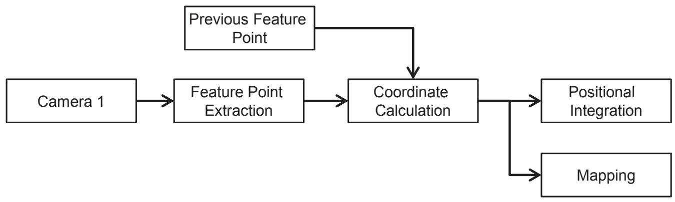

Fig. 6　SLAM Structure by Motion Stereo Method

2.2 Synchronizing Issues between Cameras

If self-position estimation is performed, SLAM using the motion stereo method is used due to the fact that Multi Angle Vision™ camera is not a stereo camera. Since it is self-position estimation technology obtaining self-position based on the image taken by the camera, it is important to grasp when the image was obtained.

A camera generally takes 30 images per one second. This is indicated in fps (frame per second) as the number of frames (frame rate) per second. The camera transmits the images captured by the image sensing device per 1/30 second, that is, at every 33.3ms. And the shutter timing is usually controlled independently by the camera itself. Although the transmitted image signals at an arbitrary timing are captured by image-capture function of ECU, which doesn't know at which timing the image was obtained.

This makes the obtained timing unclear in the range around 33.3ms though 4 cameras are mounted on Multi Angle Vision™. Even if an image is required for operation at the same timing, the image obtained at a shutter timing original for the camera will be actually transmitted.

Self-positions indicated by each camera are different while a vehicle is moving. **(Fig. 7)** This difference is equivalent to the moving distance at 33.3ms.

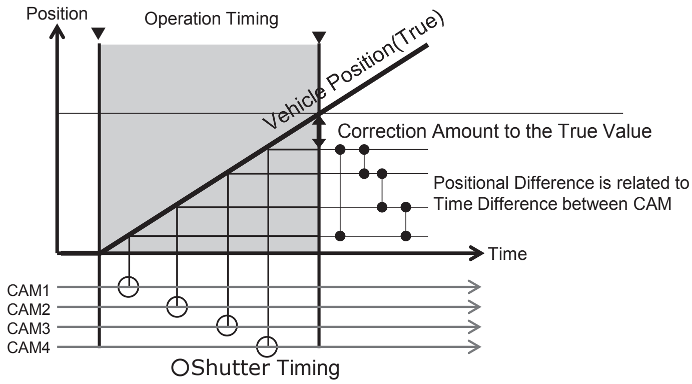

Fig. 7 Relationship between the Timing Difference and the Error within the Frame

The moving speed and error at 30fps of a single frame time are shown **(Fig. 8)**. Assuming the so-called creeping speed for an actual vehicle (5～20km), the error is 5～20cm. If the target error is ±10cm, it is a large value.

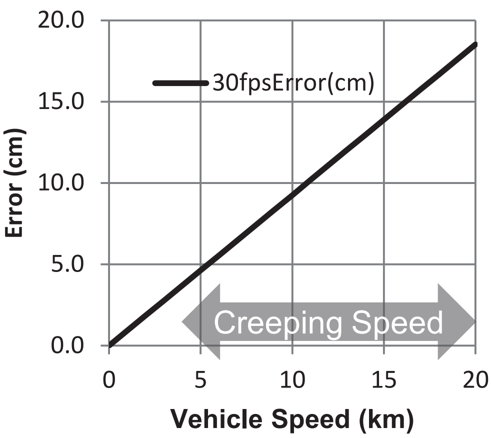

Fig. 8 Positional Error in a Single Frame

The phenomena generating the error related to the measurement timing like this may become a problem regardless of what type of sensor system periodically performing measurement is used.

It has a significant influence especially when self-position is measured in a high accuracy. Paradoxically, since Multi Angle Vision™ is originally a monitoring device and the accuracy is not so strictly required, the timing related to the image capturing is not the most important target.

This system doesn't have synchronous shutter function between cameras in the current situation.

2.3 Camera Synchronizing Method

When the shutter timing needs to be precisely controlled, hardware that enables external control of camera shutter timing is adopted. A synchronizing signal needs to be supplied separately to a camera equipped with the dedicated synchronizing function and old analog control. In addition, the command control method is used for a camera equipped with digital control. In both cases, the cost increases due to additional function such as a camera and the like.

The other method is to process and correct timing deviation in software while the hardware configuration remains unchanged.

2.3.1 Software Synchronizing Plan 1

Configuration plan to enable SLAM function to be mounted on Multi Angle Vision™ equipped with software based correction function is shown. **(Fig. 9)** Self-position coordinate is calculated from each camera image. The positional difference caused by camera timing deviation is added to this coordinate. This is the plan that the positional correction amount associated with synchronizing deviation is calculated by comparing self-position from each camera at this time and feed it back to maintain the accuracy.

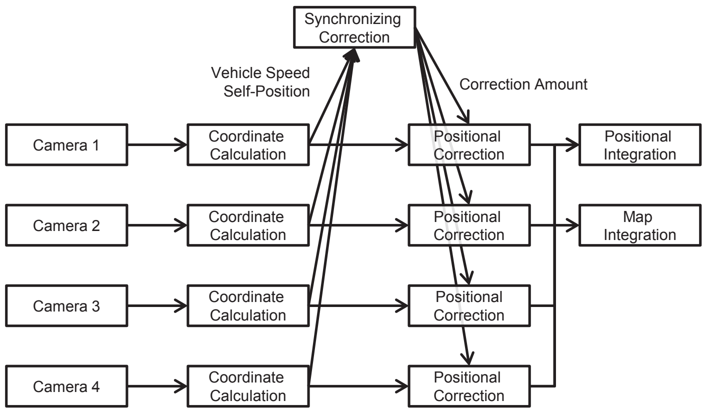

Fig. 9 Structure of Software Synchronizing Plan 1

The output of self-position in a camera timing deviation state is shown. **(Fig. 10)** The disturbance of a value associated with the timing deviation for a moving amount per one frame occurs at a moment when a vehicle starts moving. The camera timing deviation can be calculated backward by measuring this disturbance.

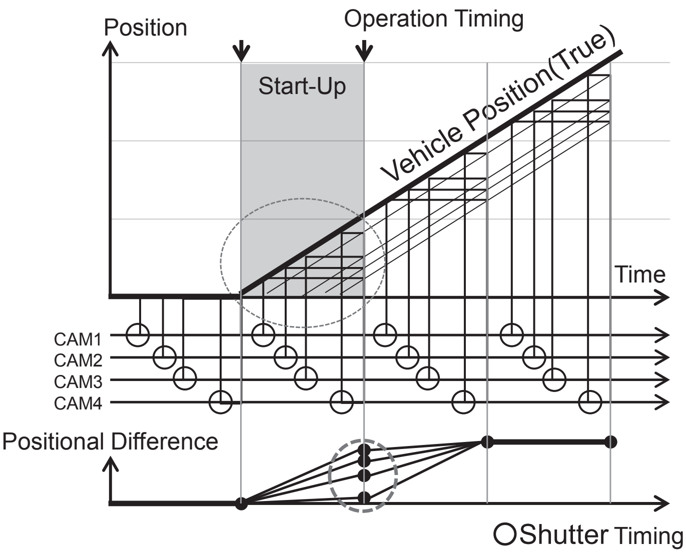

Fig. 10 Measurement Plan of Synchronizing Deviation

This plan provides the advantage that the structure is simple and independent of the types of positional measurements because the correction is performed separately from SLAM structure. On the other hand, there are the following concerns.

- Concern about an oversight of detection and a low accuracy of measurement due to only one frame after the timing capable of detecting the synchronizing deviation becomes zero speed.
- When the acceleration on start-up is low, the accuracy of synchronizing deviation is reduced due to a small positional difference.

2.3.2 Software Synchronizing Plan 2

In order to avoid the small amount of measurement that is Plan 1 concern, the direct measurement of deviation of feature point side is described in the following plan 2.

A visual field of Multi Angle Vision™ camera is illustrated from upper surface of vehicle. **(Fig. 11** left**)** The visual fields of each adjacent camera are overlapping each other. The coordinate of feature point existing in an overlapping region causes a deviation depending on the timing of each camera. Then, the feature points within the overlapping region captured by each camera are compared and the same feature points are matched. Each position difference is equivalent to synchronizing offset.

In this way, it is expected to be able to calculate at any time by operating at the feature point and reduce the error by being capable of obtaining the plural result simultaneously.

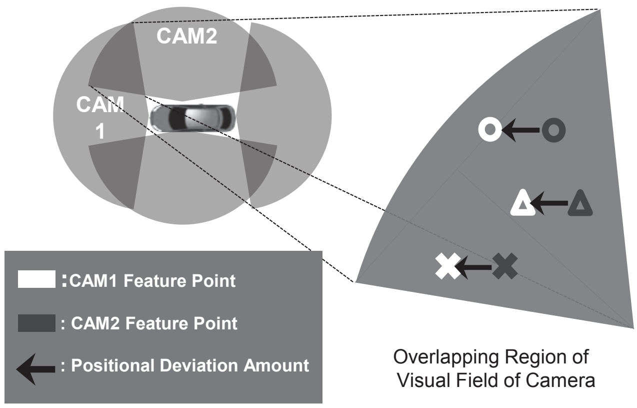

Fig. 11 Comparison of the Feature Point Coordinates in the Visual Filed Overlapping Parts

Plan 2 structure is shown in **Fig. 12**. For each camera image, a three-dimensional coordinate of the feature points is calculated first. Only an overlapping part of the results is filtered and the positional difference is obtained by synchronizing detection after finding a combination base on a feature amount. The positional correction amount is obtained from the positional difference and vehicle speed, and then the result for each camera is corrected. The result is integrated to be converted into a map and self-position.

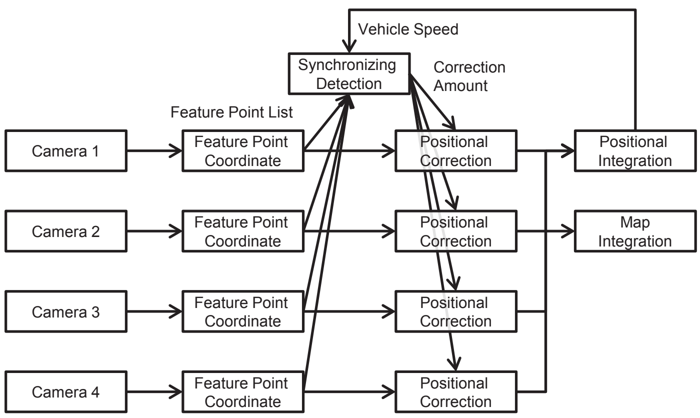

Fig. 12 Structure Software Synchronizing Plan 2

The structure becomes complex due to entering SLAM internal processing though this plan is dealing with the concerns of plan 1.

3. Conclusion

This article described the importance of the accuracy of vehicle control for the future automatic parking system. In order to deal with it, we described the self-position estimation technology to be mounted on Multi Angle Vision™, the synchronizing method between cameras which caused concern of the influence on accuracy, and two software-based methods. Although there are advantage and disadvantage with each, we would like to aim to mount the high-accuracy position estimation function on Multi Angle Vision™ by verifying these plans in future.

・Multi Angle Vision is a registered trademark of DENSO-TEN Limited.

Profiles of Writers

Wataru HASEGAWA

VICT Engineering Group Advance System R&D Dept

Toshihiro MATSUMOTO

VICT Engineering Group Advance System R&D Dept

Yuichi SUGIYAMA

VICT Engineering Group Advance System R&D Dept

Masayuki KISHIDA

VICT Engineering Group Advance System R&D Dept