Abstract—The driving support is one of the most important research areas in intelligent transport system (ITS). Moreover, obstacle detection system is one of the most important systems, and we have proposed such systems based on stereovision systems. Additionally, to assist driving safely, it is necessary to extract dynamic objects and alert drivers more quickly. In our previous reports, we proposed dynamic objects extraction methods based on Occupancy Grid Maps. However, we found that the methods suffer from disadvantages in unstable detection for dynamic objects. So, in this paper, we propose a method to analyze the motion of objects used 6D information comprising 3D position and motion of objects and extract the dynamic objects fastly based on Occupancy Grid Maps.

I. INTRODUCTION

Recently, in order to tackle such problems as traffic accidents and traffic jams, much research has been carried out into the development of Intelligent Transport System (ITS). In the field of ITS research, driving support systems are among the most important research and development fields.

To realize a driving support system, it is necessary to recognize the surrounding environment. It is well known that stereovision systems are among the most practical approaches to this problem. For example, Sekiguchi et al. [1] proposed a method to extract object information, such as preceding vehicles, lane markings, and guardrails, by using disparity images computed from individual images from two cameras. In this method, the objects are extracted by using some models to enhance the robustness. Broggi et al. [2], for the purpose of developing an autonomous vehicle, proposed a method to identify near and far obstacles by adaptively switching between three cameras. Kubota et al. [3] proposed a method to identify obstacles based on subtraction of geometrically transformed stereo images. This method has the advantage that the computational cost becomes low, because it is based on simple image subtraction. However, almost all of these methods analyze the environment using a long baseline stereovision system. For example, Sekiguchi et al. adopted a 350 mm baseline [1] and Broggi et al. used a baseline up to 1500 mm [2]. This is because a longer baseline improves the accuracy at larger distances. However, a system with an excessively long baseline introduces problems since it is difficult to be set up on a car, it is bletcherous, and it disturbs the driver’s view. On the other hand, obstacle detection methods using a monocular camera have also been proposed. Franke et al. [4], for example, suggested a method to estimate distance based on relative image motion. However, there is a problem that the distance cannot be estimated unless the ego-vehicle moves for a while. Therefore, we developed and reported on a 120 mm baseline length stereovision system [5], [6]. This offers the advantage that it can be hidden behind the rear-view mirror in the vehicle. However, this system was found to have problems in that it produced some false positives while detecting obstacles. The characteristics of such false positives were that they appeared suddenly and rarely remained in the same position for long. Such problems are likely to affect not only our system but also systems that make use of raw sensor data. Moreover, our previous system suffered from a significant deterioration in distance accuracy as the baseline was shortened. For the solution of such problems, we proposed an obstacle detection method [7] using Occupancy Grid Maps (OGM). Since, in this method, the existence of an obstacle is represented as a posterior probability based on all past measurements obtained by the stereovision system from moment to moment, the number of false positives is reduced and the deterioration in distance accuracy that arises from baseline shortening can be ameliorated.

In order to assist drivers, it is necessary to extract dynamic objects that carry a high crash risk for ego-vehicle and estimate the motion of it fastly. However, there was a problem that the proposed method requires a certain time to extract dynamic objects. This is because that, in our method, dynamic objects are extracted by accumulating an evaluation value of likelihood that the object may be a dynamic object.

So, in this paper, we propose a method to analyze the motion of objects used 6D information [8] comprised of three-dimensional position and motion of objects and extract the dynamic objects fastly based on Occupancy Grid Maps.

II. STEREOVISION SYSTEM

Figure 1(a) shows our experimental vehicle. The stereovision system in the vehicle is mounted in front of the rear-view mirror as shown in figure 1(b). The length of the baseline is about 120
mm, so it is fairly unobtrusive and is almost hidden from the driver.

The stereovision system is a sensor that can measure distances from the vehicle to obstacles based on the principle of triangulation using two or more cameras. In this system, two cameras are installed parallel to each other. For a stereovision system with a baseline length of $b$, a focal length of $f$, and a principal point of $(c_u, c_v)$, as shown in figure 2, the three-dimensional position of a target object $P(X, Y, Z)$ determined from the disparity is given by

$$
\begin{align*}
X &= b(u_x - c_u) / d \\
Y &= b(v_x - c_v) / d \\
Z &= bf / d
\end{align*}
$$

(1)

where the subscripts $L$ and $R$ refer to the left and right cameras, respectively, and $d$ is the disparity. Thus, we can obtain the three-dimensional position of a target object by computing the correspondence point between the right and left images. Figure 3 shows a disparity image where the magnitude of the disparity for each pixel is expressed with a gray scale value. In figure 3, it can be seen that there is a large amount of disparity even for textureless regions such as the road surface. This is because, in this method, a LoG (Laplacian of Gaussian) filter [9] was adopted for image enhancement and this type of filter is well known to produce such effects. In our algorithm, obstacles are detected based on this disparity image.

III. MOVING OBJECT EXTRACTION BASED ON OCCUPANCY GRID MAPS

In our previous reports, we proposed a moving object extraction algorithm based on Occupancy Grid Maps (OGM). In this section, we briefly explain the method.

A. Obstacle detection from disparity image

At first, we discuss an obstacle detection method from the disparity image. Since the disparity image has three-dimensional information on each pixel, the image projected onto the $u-d$ plane contains information concerning obstacle maps. However, an image simply projected in this manner includes not only obstacles touching the road surface but also objects above the road, such as bridges, traffic signs, and signals. Therefore, in our method, only disparity image space points satisfying $H_v^{\text{min}} < \Delta H_v < H_v^{\text{max}}$ are projected onto the $u-d$ plane, where $H_v$ is the height above the road surface, and $H_v^{\text{min}}$ and $H_v^{\text{max}}$ are predefined minimum and maximum heights above the road surface in real three-dimensional space. In our implementation, these values are

![Fig. 1 Experimental vehicle](image1)

![Fig. 2 Geometry of stereovision system](image2)

![Fig. 3 Example of disparity image](image3)

![Fig. 4 Example of $u-d$ plane projection](image4)

![Fig. 5 Example of obstacle position calculation](image5)
set to \(-0.5\) m and 2.5 m, respectively. Figure 4(b) shows an example of applying this type of projection to the image in figure 4(a).

Moreover, from \(u-d\) plane image, by extracting disparity, which has peak value on each \(u\)-line, obstacle positions can be calculated. Figure 5(a) shows an obstacle detection result. Here, obstacles are highlighted from ground level up to a constant height based on the calculated obstacle position. It can be seen from figure 5(a) that there is a false positive in an area where no object exists on the road surface, though almost all obstacles are detected accurately at their original positions. In our previous paper, we proposed a method to overcome such problems using OGM.

B. Obstacle detection using Occupancy Grid Maps

The OGM is a method to estimate posterior probability that an object exists based on all past measurements. In the OGM, the Cartesian coordinate space is partitioned into finitely many grid cells, and the posterior probability \(p(m_{x,y} | z_{1:t}, x_{1:t})\) of an obstacle being present is calculated for each cell \(m_{x,y}\) using information about the vehicle pose \(x_{1:t}\) and measurements \(z_{1:t}\) from the stereovision system. Here, \(x_{1:t}\) represents the path of the vehicle defined by the sequence of all poses up to time \(t\). Moreover, \(z_{1:t}\) denotes the set of all measurements up to time \(t\). If we assume that the surrounding objects are static, \(p(m_{x,y} | x_{1:t}, z_{1:t})\) can be recursively calculated using a Binary Bayes Filter, where the log-odds \(I_{x,y}^z\) are calculated, instead of \(p(m_{x,y} | x_{1:t}, z_{1:t})\), by

\[
I_{x,y}^z = \log \left( \frac{p(m_{x,y} | z, x_{1:t})}{1 - p(m_{x,y} | z, x_{1:t})} \right) + I_{x,y}^z
\]

(2)

where \(p(m_{x,y} | z, x_{1:t})\) is a conditional probability given by the vehicle pose \(x_{1:t}\) and a measurement \(z\) on a grid cell \(m_{x,y}\) at a time \(t\). By giving appropriate probability \(p(m_{x,y} | z, x_{1:t})\) from measurement of stereovision, which described in previous subsection, the best probabilistic obstacle map can be generated.

Figure 6 shows a comparison of obstacle detection with and without OGM. In figure 6(a) and (b), the detected obstacle positions are displayed in the original image by superimposing the raw data and the OGM. In figure 6(a), the high probability regions in OGM are displayed on the original image. Figure 6(c) shows an OGM, this figure shows that the darker the grid cell is, the higher the probability is. Mid-tone areas represent probabilities of about 0.5, where it is unknown whether an obstacle exists or not. It can be seen from these figures that some false positives are successively eliminated.

C. Moving object extraction using OGM

As mentioned above, the OGM is a method to estimate posterior probability that an object exists based on all past measurements. Therefore, if a new measurement at present time comes from stationary object, the measurement will fall on a high probability grid cell in the OGM. On the other hand, if the measurement comes from moving object, it is considered that the measurement will appear on comparatively low probability region in the OGM. Therefore, moving obstacle region can be extracted by extracting measurements, which fall on low probability region in the OGM. Each measurement is assigned a moving state \(m_{\text{State}}\) by following equation based on this probability corresponding to the measurement.

\[
m_{\text{State}} = \begin{cases} \text{Dynamic} & (0.0 \leq P < 0.3) \\ \text{Undecided} & (0.3 \leq P < 0.7) \\ \text{Static} & (0.7 \leq P \leq 1.0) \end{cases}
\]

(3)

Additionally, measurements are clustered into separate groups and each group represents a single object. In fact, measurements that distances between them are close can be regarded as a same object. Here, Each object can be assigned an object moving state \(O_{\text{State}}\) based on states of its including measurements \(m_{\text{State}}\). This object moving state \(O_{\text{State}}\) is calculated by

\[
O_{\text{State}} = \begin{cases} \text{Dynamic} & \left( \frac{N_p}{N} > T_p \right) \\ \text{Static} & \left( \frac{N_s}{N} > T_s \right) \\ \text{Undecided} & \text{otherwise} \end{cases}
\]

(4)

where \(N\) is the number of measurements in each object, \(N_p\) and \(N_s\) are the number of measurements whose moving state are Dynamic and Static, and \(T_p\) and \(T_s\) are thresholds in each case. Each object can be assigned a moving state by this method.

Figure 7 shows processing result by using above method. In figure (b), color points show moving state \(m_{\text{State}}\) of measurements, and blue, yellow and red color show Dynamic, Undecided and Static state, respectively. Additionally, in figure 7 (c), rectangles
show result of clustering into separate group and its color shows object moving state $O_{state}$ as same as $m_{state}$.

As shown in figure 7, a coming near object like an oncoming vehicle can be detected as a moving object because it appears low probability area from its existing area before. However, a preceding vehicle can’t be detected as Dynamic but as Undecided. It reasons that this object appears unknown area because it goes away from its area existing before. Here, in OGM, probability is high if object is measured sequentially at the same position. Thereby, if state of object is Undecided sequentially, it can be said its state may be Dynamic. For the reason above-mentioned, time-series processing is considered of value for the detection of moving object method.

In this paper, each object is assigned evaluation value $M_t$ to recognize moving object at time $t$. This value is calculated by following equation through time.

$$M_t = \begin{cases} M_{t-1} + 2 & (O_{state} = Dynamic) \\ M_{t-1} + 1 & (O_{state} = Undecided) \\ M_{t-1} - 2 & (O_{state} = Static) \end{cases} \quad (5)$$

If this value is over a threshold, an object that has its value is recognized as Dynamic. Also, to calculate evaluation value $M_t$ through time, it need to track object between frames. In our implementation, a simple object tracking scheme using Global Nearest Neighbor (GNN) [10] is employed to tracking method.

In addition, as shown in figure 7, the high probability region appears in the back of traveling direction of moving object. This disappearance of the region requires time, so following moving object is detected as Static if it goes into this region. Accordingly, this region needs to be initialized when moving object generating it is recognized as Dynamic. So in our method, it check if there are grids affected measurements recognized as Dynamic, and subsequently these obtained grids are initialized that probabilities are 0.5.

IV. FAST DYNAMIC OBJECT EXTRACTION USING 6D INFORMATION

As mentioned above, we described dynamic object extraction method using OGM. Figure 8(a) shows an example of extracting dynamic objects using OGM. As you can see that pedestrian and preceding vehicles are recognized as Dynamic. On the other hand, as shown in figure 8(b), there is a problem that it is difficult to extract ones fastly in the case of complex environment such that there are some follow-on vehicles in close range with each other. This is because measurements by our short baseline stereovision system are less accurate. Therefore, the follow-on vehicles on the opposite lane never fall on low probability region, since back region of most nearest vehicle on opposite lane have high probability. Hereby the evaluation value of follow-on vehicle hard to reach the threshold and, as a result, it is difficult to extract dynamic objects fastly.

For the solution of such problems, we attempt to extract dynamic objects by analyzing these motions. So, in this section, we propose a method to extract dynamic objects fastly based on 6D information, which estimate both three-dimensional position and velocity of objects.

A. 6D information based on stereovision and optical flow

Our stereovision system can measure three-dimensional position of objects as described in section II, but motion of objects can’t measure by only stereovision. To acquire the motion of objects, 6D vision algorithm fuse stereovision and
optical flow. Optical flow is computed by searching similar pixels between consecutive frames taken one camera and shows two-dimentional motion of obstacles in image. However computing optical flow for all pixels causes the long processing time. In our algorithm Kanade-Lucas-Tomasi (KLT) Tracker[11] is adapted. This tracker computes it for only feature points that can be homologized easily. Herewith, it enables to compute high-accuracy optical flow in less time. KLT-Tracker brings in two-dimensional motion of feature points between consecutive frames. Additionally, three-dimensional position of each feature points is measured by stereovision that is already developed. So fusion of optical flow and stereovision enables to acquire three-dimensional motion for each feature.

KLT-Tracker brings in two-dimensional motion of feature points between consecutive frames. Additionally, three-dimensional position of each feature points is measured by stereovision that is already developed. So fusion of optical flow and stereovision enables to acquire three-dimensional motion for each feature.

However, this 6D information is astable information by only this fusion, because of less-accurate measured three-dimensional position. So, to acquire more stable object motion, this 6D vision algorithm fuses stereovision and optical flow by Kalman filter. Applying this filter can estimate more robust three-dimensional position and additionally, reliable velocity at the same time.

Figure 9(a) and (b) shows the result of estimation 6D information for each features selected KLT-Tracker. In these figures, features recognized as Dynamic are described as purple points, and additionally line lengthens behind a point shows motion vector. Briefly, the more line is long, the more the feature is estimated fast-moving. On the other hand, recognized as Static are described as green points.

**B. Association of 6D information with OGM**

As noted on previous section, three-dimensional motion of objects can be analyzed. So, in our method, motion of each object is analyzed by 6D information and evaluation value shown equation (5) is recalculated.

First, 6D information of feature points is associated with the bounding box clustered measurements, and then the motion of objects are analyzed by checking a state of feature points in the bounding box. This state is evaluated by following equation and a moving state $F_{state}$ is assigned to each object,

$$
F_{State} = \begin{cases} 
\text{Dynamic} & (F_D / F > S_D) \\
\text{Static} & (F_S / F > S_S) \\
\text{Undecided} & \text{otherwise}
\end{cases}
$$

where $F$ is total number of feature points in each object, $F_D$ and $F_S$ are the number of feature points whose moving state are Dynamic or Static, and $S_D$ and $S_S$ are thresholds in each case. Each object can be assigned a moving state by this method.

However, there was a problem that 6D information of feature
points aren’t estimated correctly, because of less-accurate measurements by our short baseline stereovision system. For the solution of the problem, time-series processing of value for the extraction of dynamic objects method is again considered. Therefore, evaluation value $M_t$ of each object is recalculated by following equation through time.

$$M_t = \begin{cases} M_{t-1} + 2 & (F_{state} = \text{Dynamic}) \\ M_{t-1} & (F_{state} = \text{Undecided}) \\ M_{t-1} - 2 & (F_{state} = \text{Static}) \end{cases}$$ (7)

If this value exceeds a certain threshold, the object is recognized as Dynamic.

V. EXPERIMENT

As noted above, we described the method of detecting objects and extracting dynamic objects by OGM, and additionally we propose the fast dynamic objects extraction method by 6D information. In this chapter, we present the results of the method and evaluate its validity.

The results of dynamic objects extraction using the proposed method (fusion of 6D information) are compared with the those using the previous method (only OGM). Figure 10 shows the complex urban scene that dynamic objects are extracted difficultly by previous method. It can be seen from figure 10(a) that some objects don’t extract as Dynamic, for example follow-on vehicles in opposite lane. This is because the follow-on vehicle goes into high priority region affected by preceding vehicle. On the other hand, as shown in figure 10(b), it can be seen that the bounding box of follow-on vehicle contains much feature points evaluated as Dynamic, and for this reason, this objects are evaluated as Dynamic objects from the point of view of 6D information.

Then, figure 11 shows comparison of the dynamic object extraction distance by proposed method and previous method in the scene that some vehicles are oncoming in sequence. The number of this figure shows turn that an oncoming car comes close. It can be seen from this figure that our proposed method can extract dynamic ones more father distance than previous method. Especially, as shown figure 12(a), the tenth vehicle isn’t evaluated as dynamic by previous method, because the vehicle is occlusion by the truck runs in the front and isn’t evaluated as dynamic. On the other hand, in proposed method, it can be seen from figure 12(b) that feature points on the vehicle are estimated as dynamic and, for the result, the vehicle is extracted as dynamic objects. Hereby, in complex scene where many objects exist, our proposed method can extract dynamic objects in early stage.

VI. CONCLUSION

In this paper, we describe the method of detection objects and extraction dynamic objects by OGM, and additionally we propose the fast dynamic objects extraction method by 6D information. Our method is summarized as follows:

- By using OGM that deal with all measurements through time, false detection could be deleted and dense obstacle map could be generated.
- By using 6D information comprised of three-dimensional position and motion of objects, in complex urban scene, dynamic objects extract robustly in early stage.

In view of these results, it can be concluded that associating 6D information with objects detected using OGM is an effective dynamic objects extraction method in complex scene. For the future work, we want to develop an automatic driving vehicle with stereovision system. For the realization of this, we are slated for speeding up of processing time of our system.

REFERENCES