A Comparative Study on Ground Surface Reconstruction for Rough Terrain Exploration

K. Otsu*, M. Otsuki**, T. Kubota**

*Department of Electric Engineering and Information Systems, The University of Tokyo, Japan
e-mail: kyon@ac.jaxa.jp

**Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency, Japan
e-mail: otsuki.masatsugu@jaxa.jp, kubota@isas.jaxa.jp

Abstract

Obtaining information about foreseeing terrain is an important technique for mobile robot navigation in unstructured terrain. There are several methods proposed in the literature which try to extract the geometrical models of ground surfaces efficiently. This paper categorizes the conventional ground approximation methods into three types, and presents the conceptual and performance differences thorough synthetic and real data analysis. The results demonstrate the effectiveness of previous and newly developed techniques, and also provide discussions which will be useful in the design of terrain modeling as a step of mobile robot navigation.

1 Introduction

Some of the important applications of robotics exist in unstructured fields such as agriculture, mining, search and rescue, defense, collapsed buildings, and space. Especially, robots are expected to perform tasks in hazardous areas in behalf of human beings. In such ultimate environments, robots should have autonomous capabilities to keep the safety from hazardous properties, e.g., rocks, ditches, and steep slopes. Furthermore, onboard autonomy enables a robot to perform without human confirmation, and thereby enhances efficiency in mission operations. Capabilities for autonomous mobile robots include environmental perception, assessment of traversability, local and global path generation, and pose estimation.

Among all challenges to be overcome, this paper addresses the problem of perception and terrain modeling in order to enable safe and reliable operations in unstructured natural terrain. Objectives are detecting and locating terrain features that are hazardous by mobility constraints. Mobility hazards include step obstacles, extreme slopes, deep ditches, cliffs, and all other features that spoil robot's safety. These geometric hazards can be detected with exteroceptive look-ahead sensors such as passive stereo cameras and active LIDAR (light detection and ranging) sensors. Those sensors provide micro-properties about surrounding terrain, which should be converted to macro-properties considering vehicles' dimensions and mobility systems. The generated maps are often called digital elevation maps (DEMs) which are grid-based representations containing terrain elevations (Fig. 1).

Several groups have been tackled with this problem. One of the most popular works has been done by Mars rover teams in NASA/JPL [1, 2, 3]. They use stereo cameras to generate a cloud of 3D points and convert it to a DEM within GESTALT algorithm [4]. LIDARs can also be used for planetary terrain mapping [5, 6, 7]. Important techniques are developed in other fields such as intelligent automobiles [8, 9] and automated agriculture [10].

There have been different approaches to approximate ground surfaces with different fields and assumptions. This paper provides a comparative study on those approaches from conceptual differences, without going into the detail of specific implementations. The main objective is to understand the difference and valid conditions for distinct approaches, and provide a guidance to the development of autonomous perception systems.

2 Stereo Vision

In this paper, a passive stereo vision system is basically used. Besides the mechanical and electrical advantages, it provides rich information as arrays of RGB triplets. This section briefly reviews the geometry of stereo vision for recovering 3D point clouds from images.

2.1 Geometrical model

A standard configuration of stereo cameras is shown in Fig. 2. The cameras are set horizontally so that their optical axes are parallelized. Parameters are the baseline between cameras (b), the camera height from ground (h), and the tilt angle (θ). The origin of the camera coordinate system is set to the optical center of the left camera. Images are rectified so that epipolar lines are horizontal.

An image coordinate systems uvd is defined by the position of projected world point (u_c, v_c) and the relative disparity between left and right images d. Assuming the

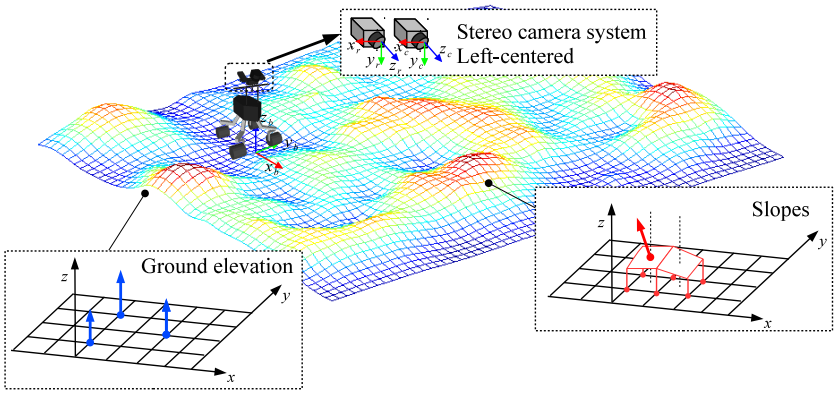


Figure 1. : Terrain modeling and hazard detection for ground vehicles

pin-hole camera model, a world point $P_c = (x_c, y_c, z_c)$ in the camera coordinate system is projected on a image plane by

$$\begin{bmatrix} u_c \\ v_c \\ d \end{bmatrix} = \frac{f}{z_c} \begin{bmatrix} x_c \\ y_c \\ b \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \\ 0 \end{bmatrix} \tag{1}$$

where f is the focal length (assuming f is proportional both in u and v direction), and (u_0, v_0) is the projection of the optical center. Through these equations, a point can be converted bidirectionally between the image and world coordinate systems. Later in this paper, the subscript $(\cdot)_c$ is omitted to simplify the notation.

2.2 Correspondence problem

A significant part of stereo vision is finding point correspondences from left and right images. This corresponding problem is regarded as a challenging problem especially in outdoor systems due to noise, ill illumination, and poor textures. The methods for searching correspondences can be distinguished into two categories:

- *Area-based*: these algorithms compute the correlation of a small patch (window) in one image to another. Popular methods for correlation are the Sum of Absolute Differences (SAD), the Sum of Squared Differences (SSD), and the Normalized Cross Correlation (NCC). Area-based methods produces dense disparity maps as shown in Fig. 3.
- *Feature-based*: these algorithms extract distinct features (e.g., corners) from the images and find the correspondences between extracted features. A number of salient feature detectors have been studied including corner detectors (e.g., [13]), edge detectors (e.g., [14]), and blob detectors (e.g., [15]). Feature-based methods are faster and more robust while they provide only sparse depth information.

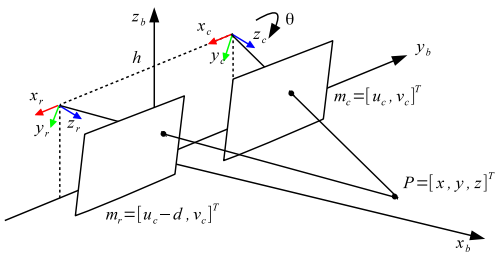


Figure 2. : Stereo geometry

Most methods for ground surface reconstruction require dense depth information. Therefore, area-based methods are typically used for rough terrain modeling. However, area-based methods are prone to wrong correspondences. Recently, several optimization techniques based on energy minimization have been developed which incorporate smoothness in neighboring pixels (e.g., [12]). The energy to be minimized is typically defined as

$$\mathcal{E}(\mathcal{X}) = \lambda_D \mathcal{E}_D(\mathcal{X}) + \lambda_S \mathcal{E}_S(\mathcal{X}) \tag{2}$$

for a configuration of disparities $\mathcal{X}$. $\mathcal{E}(\mathcal{X})$ is decomposed into a linear combination of two components: $\mathcal{E}_D(\mathcal{X})$, a measure of disagreement between $\mathcal{X}$ and observed data, and $\mathcal{E}_S(\mathcal{X})$, piecewise smoothness.

Energy functions of the form (2) can arise with Markov Random Fields (MRFs) formulation. The MRFs are sets of random variables such that a variable is conditionally dependent only on its neighbors. The reconstruction problem can be considered as maximum a priori (MAP) estimation given an observation $\mathcal{D}$, that is,

$$\mathcal{X}^* = \arg\max_{\mathcal{X}} P(\mathcal{X}|\mathcal{D}) \tag{3}$$

$$= \arg\max_{\mathcal{X}} P(\mathcal{D}|\mathcal{X})P(\mathcal{X}) \tag{4}$$

$$\because P(\mathcal{X}|\mathcal{D}) = P(\mathcal{D}|\mathcal{X})P(\mathcal{X})/P(\mathcal{D}) \tag{5}$$

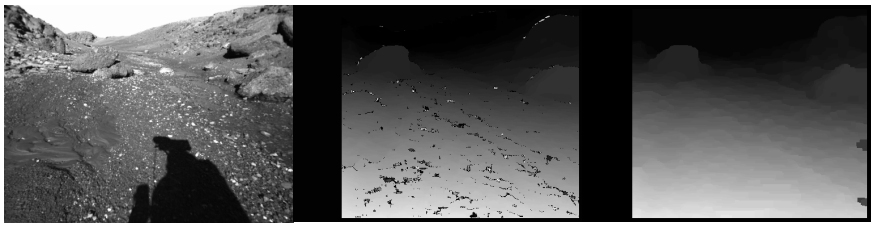


Figure 3. : SAD-based disparity maps computed from stereo images in dataset [11]. The standard area-based correlation (center) contains speckle noises, while they are eliminated with optimization using graph cuts (right).

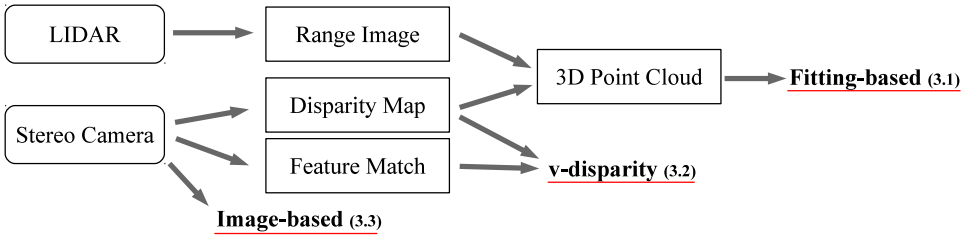


Figure 4. : Three ground reconstruction categories and processing flows from sensors

where $P(\mathcal{D}|\mathcal{X})$ and $P(\mathcal{X})$ are corresponding to $\mathcal{E}_D(\mathcal{X})$ and $\mathcal{E}_S(\mathcal{X})$ respectively. The MAP estimation is generally NP-hard, therefore it is solved with efficient approximation methods such as graph cuts [12]. The disparity maps computed with SAD and graph cuts are shown in Fig. 3. The optimization technique provides a good disparity map, while computational burden should be considered (time differs by 2–3 orders of magnitude).

3 Ground Surface Reconstruction

The ground surfaces are reconstructed from the measurements using stereo cameras or LIDARs. Different algorithms have been proposed in the literature for rough terrain modeling. In this paper, those methods are categorized into three types in terms of data they require. The relation is illustrated in Fig. 4.

3.1 From point cloud

A fitting-based method using a point cloud is probably the most common way to approximate terrain in rough terrain [4, 16, 17, 18, 19, 9]. A set of data points represented in the Euclidean 3D space (x, y, z) are fitted by (weighted) least squares to planes [16, 17, 19] or arbitrary algebraic forms of ground models [9]. A plane $z = ax + by + c$ is often used to represent a small patch of an uneven surface around (x, y). Variants employ different weighting rules to each data points, such as sensor uncertainty [16] and residuals from fitted planes [17, 19].

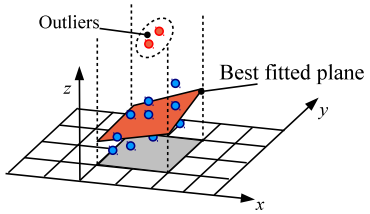


Figure 5. : Plane fitting from point cloud with outliers

A crucial problem for this approach is handling outliers. As shown in Fig. 5, 3D point measurements may contain outliers caused by sensor noise and wrong correspondences. Since least-squares fitting is sensitive to errors, it should be removed using robust techniques such as RANSAC [20] or incremental weighting [17].

The advantages of the fitting-based approach are:

- The ground structure is well represented if sufficient resolution is selected.
- The map registration is relatively easy.

On the other hand, the disadvantages are:

- A dense point cloud is required.
- Outliers should be removed.
- Sensor uncertainty is hard to be transformed to elevation uncertainty (further discussion found in [21]).

3.2 From v-Disparity representation

The *v*-disparity approach was first introduced in the field of intelligent automobiles [8], and developed for on-load and off-load vehicles [22, 23, 24, 25] . It estimates approximated road surface without explicit 3D point computation. The required assumption is that large areas of the road consist of planes.

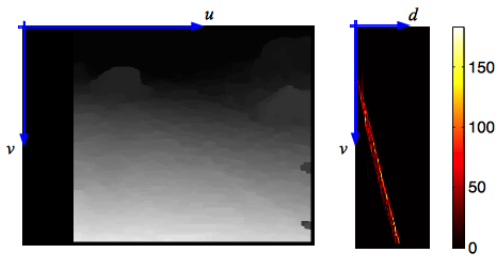


Figure 6. : Disparity and v-disparity space representation for the image in Fig. 3. The ground surface forms a straight line in v-disparity space.

Suppose that the ground surface is tilted by Θ with respect to camera's optical axis ($\Theta = \theta$ for completely planar terrain). The ground surface is modeled as

$$z = -\frac{1}{\tan\Theta}y + \frac{h}{\sin\Theta} \tag{6}$$

using a simple geometry. Substituting (6) with (1) yields

$$v = \frac{h}{b\cos\Theta}d - f\tan\Theta + v_0 \tag{7}$$

$$\propto d \tag{8}$$

which suggests that the plane in the world coordinate systems is projected along a straight line in the v-disparity(d) space. Fig. 6 shows the v-disparity notation of the image in Fig. 3. In this example, a dense disparity map is converted to the v-disparity space using a voting scheme.

Then, the ground planes are extracted from the v-disparity image. In the literature, several methods can be found such as the Hough transform [8, 23, 24], candidate line fitting [22], and statistical energy computation [25]. Except [25], segments of straight lines are fitted in the v-disparity image. However, such assumption is only valid if the most of grounds is composed of planes. To adapt the v-disparity approach to complex rough terrain, modification of the algorithm is needed.

This paper formulates an efficient method based on dynamic programming. The method is briefly depicted in Fig. 7. Starting from a v-disparity image, the vote scores are propagated to succeeding rows by the update rule

$$a_{i,j}^{new} = a_{i,j}^{old} + \max(a_{i-1,j-1}\ ,\ a_{i-1,j}\ ,\ a_{i-1,j+1}) \tag{9}$$

where $a_{i,j}$ denotes the value at (i, j) in the v-disparity map. After reaching the bottom row, the image is backtracked so that the values on a path are greater than the others in the same rows. The selected path, which consist of both straight line segments and curves, is the expected ground surface. Fig. 8 shows the results for two types of terrain (flat and undulating).

Furthermore, to generate a DEM from the representations, a disparity image is vertically divided based on the geometry in Fig. 2. This division allows the computation of ground surfaces along each y_b.

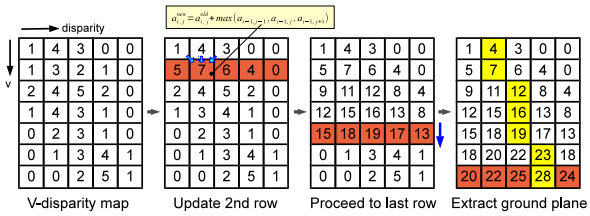
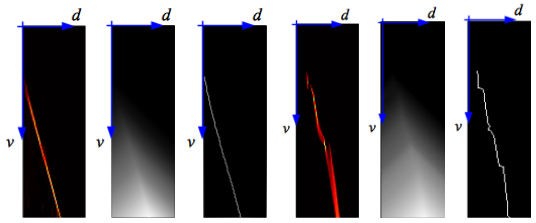


Figure 7. : A brief description of ground surface extraction algorithm using dynamic programming



(a) Flat plane (b) Undulating ground

Figure 8. : The ground estimation results on v-disparity images (Left to right: v-disparity image, accumulated score, and extracted main ground)

The advantages of the v-disparity approach are:

- The explicit 3D point computation can be avoided.
- The disparity map can be either dense or sparse.
- The voting scheme helps outlier removal.

On the other hand, the disadvantages are:

- The capability of terrain representation is limited.
- The view should be mostly covered with planes.

3.3 From image

A surface structure can be directly recovered from images. This technique was intensively studied in the field of photogrammetry and computer graphics [26, 27, 28]. The basic idea is deforming a surface mesh so that it is consistent with input images. Traditionally, triangular meshes are used to model the surfaces, and the position of vertices are controlled to minimize an objective function. This paper employs the 2.5D notation of $z_b = f(x_b, y_b)$ for every grid point of the DEM, and searches the best configuration of z_b that minimizes a cost function.

As shown in Fig. 9, the projection of a vertical line at (x_b, y_b) forms a line on an image. The idea is finding a point that has the strongest correlation with the corresponding point on the other image. Hence, this method has the advantage in computational cost since it requires only a small search space for finding matches. The correlations using various metrics are computed in Fig. 10.

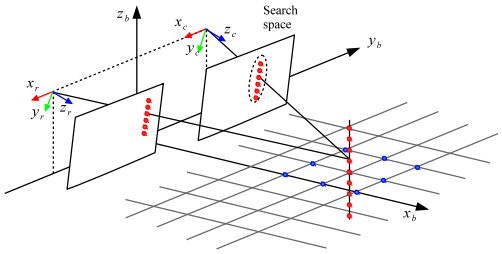


Figure 9. : Projection of vertical line in world coordinates to image planes

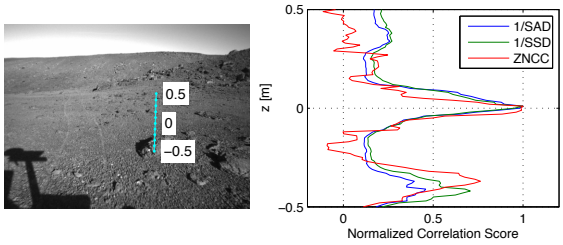


Figure 10. : Correlation scores for different elevation by 1[cm] intervals. The vertical line at $(3.0, -0.5)$ is shown in the left image. Two correlation peaks are observed; one is correct (around $z = 0$[m]) and the other is spurious (around $z = -0.4$[m]).

Two correlation peaks are observed at $z = 0$ and -0.4[m]. (In fact, the former is correct.)

To avoid local minima, smoothness constraint is considered similar to (2). The energy function in this paper is organized as follows. For a DEM configuration $\mathcal{X}$, the data term $\mathcal{E}_D(\mathcal{X})$ is defined as the sum of SAD

$$\mathcal{E}_D(\mathcal{X}) = \sum_{\mathbf{x}\in\mathcal{X}} \left[\sum_{\mathbf{u}\in C_c(\mathbf{x}), \mathbf{v}\in C_r(\mathbf{x})} |I_c(\mathbf{u}) - I_r(\mathbf{v})| \right] \quad (10)$$

where $C_{(\cdot)}(\mathbf{x})$ is a small image patch whose center is the projection of a world point $\mathbf{x}$, and $I_{(\cdot)}(\mathbf{u})$ is the intensity value at $\mathbf{u}$. The smoothness term is defined as the elevation gaps between neighboring points

$$\mathcal{E}_S(\mathcal{X}) = \sum_{\mathbf{x},\mathbf{y}\in\mathcal{N}} [\min(g, |\mathbf{x}_z - \mathbf{y}_z|)] \quad (11)$$

where $\mathcal{N}$ is a set of neighboring points, and g is a constant. The optimization problem using this function can be solved efficiently with the α-expansion algorithm [12].

The advantages of the image-based approach are:

- Only sparse matching are required.
- The search space for matching is limited along a short line on the image.

On the other hand, the disadvantages are:

- The method may be vulnerable to wrong matches.
- Occluded spaces should be handled.

3.4 Other reconstruction methods

Other methods should be briefly described. There is the Locus algorithm [21] which finds the intersection of vertical line projections and sensor measurements in the depth profile space. A similar method is used in [29] which computes the intersection of projected vertical lines for stereo camera.

Apart from ordinary grid-based DEMs, other terrain representations can be considered. An acquired point cloud can be converted to a irregular triangular mesh [5, 6], which supports variable resolution and complex shapes. A DEM can be considered not in cartesian coordinates but in cylindrical coordinates. The Cylindrical Coordinate DEM (C^2DEM) is proposed in [7] which handles the variable density of sensor measurements properly.

4 Experimental Results

The behavior of three different methods are compared using synthetic data. Later, the performance is shown for ground surface reconstruction using real rover dataset.

4.1 Synthetic data

To demonstrate the properties of three ground approximation methods, an analysis using synthetic data is conducted. The ground data points are generated in a 5×6 [m] range with topographical relief modeled by Gaussian distributions. A sample scene generated is shown in Fig. 11 and simulation parameters are shown in Table 1. The ground points are projected onto a stereo camera which is located at $(0, 0, 1)$ and tilted downward by 20 [deg]. The ground-truth disparity map is calculated from the synthetic projection.

The accuracy of ground approximation is evaluated by the RMS error between approximated planar segments and 3D points, which is computed by

$$E_{RMS} = \sqrt{\frac{\sum_{i=1}^{M_g} \sum_{j=1}^{N_i} (z_{ij} - a_i x_{ij} - b_i y_{ij} - c_i)^2}{\sum_{i=1}^{M_g} N_i}} \quad (12)$$

where M_g is the number of grids and N_i is the number of data points in i-th grid. The planar parameters are computed by least squares using four neighborhoods.

The bottom figures in Fig. 11 show the reconstruction results using three methods. In these figures, only elevation on grids are shown. This experiment assumes the ideal conditions; no noise or wrong matches are considered. The fitting-based method (left) exhibits an accurate ground approximation, while there are some holes which should be interpolate to obtain full terrain knowledge. The v-disparity approach (center) provides a less reliable map for uneven terrain. It is straightforward since the v-disparity ground extraction relies on the existence

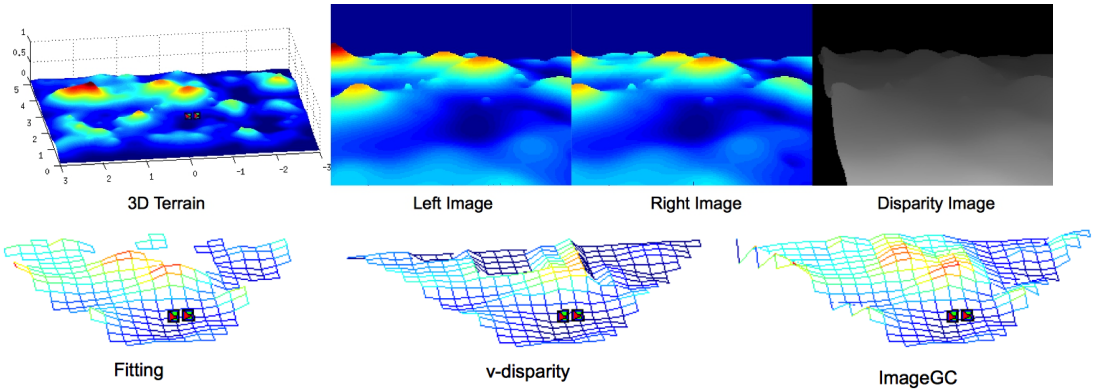


Figure 11. : Synthetic data example (top row) and ground approximation results by fitting, *v*-disparity, and image-based methods (bottom row)

Table 1. : Simulation parameters

Parameter	Default
Camera parameters	
Focal length f	240 [pixel]
Image resolution	256x192 [pixel]
Optical center (u_0, v_0)	Center of image
Camera height h	1.0 [m]
Camera tilt θ	20 [deg]
Baseline b	0.20 [m]
Terrain parameters	
Data range	x_b: 0 ~ 5 [m]
	y_b: −3 ~ 3 [m]
Data resolution	0.005 [m]
Topographical relief	Sum of Gaussians
Disparity noise σ_d	Gaussian (0 ~ 5 [pixel])
Speckle noise ρ_s	White (0 ~ 20[%])
DEM parameters	
Grid size	0.10, 0.20 [m]

of dominant planes. The image-based method (right) also describes an adequate trend of the terrain, although this experiment assumes perfect feature correspondences.

Next, the influence of noise is quantitatively studied using Monte Carlo simulation. The following noise are artificially injected to synthetic disparity images:

- *Pixel localization uncertainty*: the Gaussian noise with deviation σ_d is injected to disparity values. It derives from image sampling and improper localization of corresponding points.
- *Wrong matches*: the white noise randomly appeared on a dense disparity image is injected with a ratio ρ_s of whole image. Such speckle noise can be caused by wrong correspondences as shown in Fig. 3.

Fig. 12 shows the impact of pixel uncertainty. The fitting-based method shows the highest stability for pixel noise. The image-based method is as accurate as the

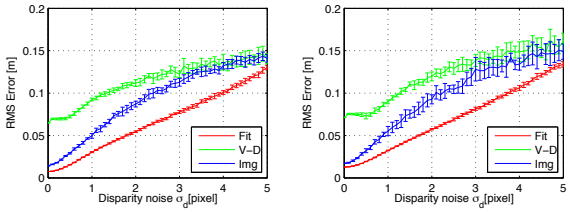


Figure 12. : Ground accuracy vs pixel noise (grid size:left 0.10[m], right 0.20[m]). The fitting-based method shows the highest stability against pixel measurement error.

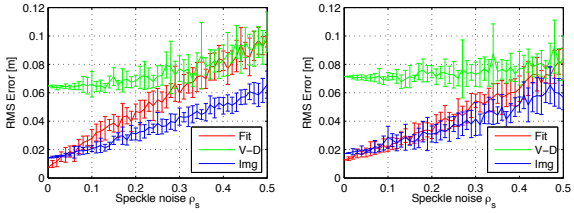
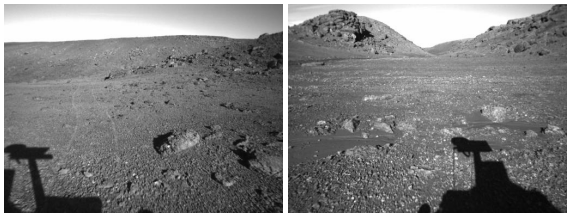


Figure 13. : Ground accuracy vs speckle noise rate (grid size: left 0.10[m], right 0.20[m]). The image-based method is the most stable against speckle noise caused by wrong matches.

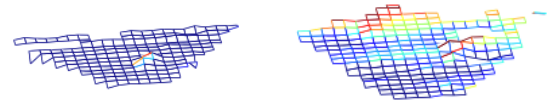
fitting-based method if noise is negligible, while its accuracy and consistency are degraded in function of noise magnitude. On the other hand, the speckle noise stability in Fig. 13 shows that the image-based method is more robust than the other methods, since the image-based method only relies on the correlation along world vertical lines. The performance of plane fitting is degraded slightly worth with the presence of speckle noise. Therefore, it should be used with complete disparity maps.

4.2 Real images

The results using real images of rough terrain are shown in Fig. 14. The Devon Island rover navigation



(a) Input images



(b) Fitting-based method result

(c) v-dispairty-based method result

(d) Image-based method result

Figure 14. : Rover image results

dataset [11] is used, which captures images with Bublebee XB3 camera of 0.24[m] baseline. A 0.02[m]-grid DEM is generated using three distinct methods.

The results show the validity of three methods for real imagery. Every method provides sufficient maps for navigation. Since the fields where a rover can traverse are supposed to be dominated by planes, the v-disparity approach was able to represent ground surfaces. The Hough transform-based ground detection could also work in these examples.

5 Conclusions

This paper has categorized and evaluated the conventional ground approximation methods using exteroceptive sensors. There are three approaches that are widely used for rough terrain modeling. A numeric analysis was conducted using Monte Carlo simulation which considers pixel measurement error and stereo matching error. The results can be used as guidelines in the design of terrain modeling for robot navigation. The methods were also validated using real images from rover stereo camera in a natural field.

Future works include more comprehensive studies on those methods. For example, the simulation are to be performed for different types of terrain containing slopes, cliffs, and ditches. The impact of camera height and tilt should be considered. Furthermore, real image analysis including quantitative error analysis and timing analysis will be done on different scenes.

Understanding geometrical properties of foreground scenes is important problem for mobile robotics. The approaches in this paper have inevitable tradeoffs. Those methods should be used adaptively to terrain considering hazards, mission efficiency, and requirement from navigation systems.

Acknowledgment

This work was supported by Grant-in-Aid for JSPS Fellows.

References

[1] M. Maimone, A. Johnson, Y. Cheng, R. Willson, and L. H. Matthies, “Autonomous navigation results from the Mars Exploration Rover (MER) mission,” in *International Symposium on Experimental Robotics*, 2004.

[2] M. W. Maimone, P. C. Leger, and J. J. Biesiadecki, “Overview of the Mars Exploration Rovers’ Autonomous Mobility and Vision Capabilities,” in *International Conference on Robotics and Automation (ICRA) Space Robotics Workshop*, 2007.

[3] J. P. Grotzinger, J. Crisp, A. R. Vasavada, and Others, “Mars Science Laboratory Mission and Science Investigation,” *Space Science Reviews*, vol. 170, no. 1-4, pp. 5–56, 2012.

[4] S. Goldberg, M. Maimone, and L. Matthies, “Stereo vision and rover navigation software for planetary exploration,” in *IEEE Aerospace Conference*, vol. 5, pp. 2025–2036, 2002.

[5] E. Dupuis, P. Allard, J. Bakambu, T. Lamarche, W.-h. Zhu, and I. Rekleitis, “Toward autonomous long-range navigation,” in *International Symposium on Artificial Intelligence for Robotics and Automation in Space*, 2005.

[6] D. Gingras, T. Lamarche, J. L. Bedwani, and E. Dupuis, “Rough Terrain Reconstruction for Rover Motion Planning,” in *Canadian Conference on Computer and Robot Vision*, pp. 191–198, 2010.

[7] G. Ishigami, M. Otsuki, and T. Kubota, “Range-dependent Terrain Mapping and Multipath Planning using Cylindrical Coordinates for a Planetary Exploration Rover,” *Journal of Field Robotics*, vol. 30, no. 4, pp. 536–551, 2013.

[8] R. Labayrade, D. Aubert, and J.-P. Tarel, "Real time obstacle detection in stereovision on non flat road geometry through "v-disparity" representation," in *IEEE Intelligent Vehicle Symposium*, vol. 2, pp. 646–651, 2002.

[9] F. Oniga and S. Nedevschi, "Processing Dense Stereo Data Using Elevation Maps: Road Surface, Traffic Isle, and Obstacle Detection," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 3, pp. 1172–1182, 2010.

[10] C. Wellington, A. Courville, and A. T. Stentz, "Interacting Markov Random Fields for Simultaneous Terrain Modeling and Obstacle Detection.," in *Robotics: Science and Systems Conference*, 2005.

[11] P. Furgale, P. Carle, J. Enright, and T. D. Barfoot, "The Devon Island rover navigation dataset," *The International Journal of Robotics Research*, vol. 31, no. 6, pp. 707–713, 2012.

[12] Y. Boykov, O. Veksler, and R. Zabih, "Fast Approximate Energy Minimization via Graph Cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1–18, 2001.

[13] C. Harris and M. Stephens, "A Combined Corner and Edge Detector," in *Alvey Vision Conference*, vol. 15, pp. 147–151, 1988.

[14] J. Canny, "A computational approach to edge detection.," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, pp. 679–98, June 1986.

[15] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[16] Z. Zhang, "A stereovision system for a planetary rover: Calibration, correlation, registration, and fusion," in *IEEE Workshop on Planetary Rover Technology and Systems*, 1996.

[17] D. Gennery, "Traversability analysis and path planning for a planetary rover," *Autonomous Robots*, vol. 6, no. 2, pp. 131–146, 1999.

[18] A. Chilian and H. Hirschmuller, "Stereo camera based navigation of mobile robots on rough terrain," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4571–4576, 2009.

[19] T. Saitoh, M. Suzuki, and Y. Kuroda, "Vision-Based Probabilistic Map Estimation with an Inclined Surface Grid for Rough Terrain Rover Navigation," 2010.

[20] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[21] I. Kweon and T. Kanade, "High-Resolution Terrain Map from Multiple Sensor Data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 278–292, 1992.

[22] A. Broggi and C. Caraffi, "Obstacle detection with stereo vision for off-road vehicle navigation," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, 2005.

[23] J. Zhao, J. Katupitiya, and J. Ward, "Global Correlation Based Ground Plane Estimation Using V-Disparity Image," in *IEEE International Conference on Robotics and Automation*, pp. 529–534, 2007.

[24] J. Zhao, M. Whitty, and J. Katupitiya, "Detection of non-flat ground surfaces using V-Disparity images," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4584–4589, Oct. 2009.

[25] Y. Cong, J. Peng, J. Sun, L. Zhu, and Y. Tang, "V-disparity Based UGV Obstacle Detection in Rough Outdoor Terrain," *Acta Automatica Sinica*, vol. 36, no. 5, pp. 667–673, 2010.

[26] P. Fua and Y. G. Leclerc, "Object-centered surface reconstruction: Combining multi-image stereo and shading," *International Journal of Computer Vision*, vol. 16, no. 1, pp. 35–56, 1995.

[27] L. Zhang and M. Seitz, "Image-Based Multiresolution Modeling by Surface Deformation," tech. rep., Carnegie Mellon University, 2000.

[28] S. Sugimoto, K. Motooka, and M. Okutomi, "Direct Generation of Regular-Grid Ground Surface Map from In-Vehicle Stereo Image Sequences," in *IEEE International Conference on Computer Vision Workshops*, pp. 600–607, Ieee, 2013.

[29] M. Vergauwen, M. Pollefeys, and L. Van Gool, "A stereo-vision system for support of planetary surface exploration," *Machine Vision and Applications*, vol. 14, no. 1, pp. 5–14, 2003.