Image Feature-based Traversability Analysis for Mobile Robot Navigation in Outdoor Environment

Mohammed Abdessamad BEKHTI, Shizuoka University, bmabdessamed@gmail.com
Soudai TANAKA, Shizuoka University, f0010072@ipc.shizuoka.ac.jp
Yuichi KOBAYASHI, Shizuoka University, tykobay@ipc.shizuoka.ac.jp
Toru KANEKO, Shizuoka University, tmtkane@ipc.shizuoka.ac.jp

For an autonomous mobile robot, an important task to accomplish while maneuvering in outdoor rugged environments is terrain traversability analyzing. Due to the large variety of terrain, a general representation cannot be obtained a priori. Thus, the ability to determine the traversability based on the vehicle motion information and its environments is necessary, and more likely to enable access to interesting sites while insuring the soundness and stability of the mobile robot. We introduce a novel method which can predict motion information based on extracted image features from outdoor university campus environments, to finally estimate the traversability of terrains. A wheeled mobile robot equipped with an optical sensor and an acceleration sensor was used to conduct experiments.

Key Words: Image recognition, Mobile robot, Terrain traversability analysis

1 Introduction

Autonomous mobile robot has been investigated for more than 25 years. Scientists aimed to reduce human monitoring and maintenance by turning mobile robots into intelligent machines, able to identify features, learn from experience, build maps, take decisions based on the surrounding environment and navigate autonomously [1]. Unmanned ground vehicles have proved effective autonomous operations in miscellaneous environments configuration as desert, farms, and urban environments [2] [3] [4]. Perception of both indoor and outdoor environments differs, since they are fundamentally different in nature, thus navigation challenges differ. When navigating in rugged territories, assessing the traversability of a terrain is a key feature of any unmanned ground vehicle [5]. Different characteristics related to the terrain can be employed to depict the traversability including texture, geometry, and/or characteristics related to the vehicle, energy required to cross a terrain, or instability issues [6]. The objective is to reduce at most encountering situations where the platform's safety and stability may be endangered. In some extreme environments, such as on Mars, terrain deformation can be caused due to robot force and weight, which will influence its attitude and configuration [7]. Same case scenario can be witnessed in sandy terrains, when strong wheel slip lead the wheels to sink in the ground [8].

Terrain traversability estimation has been addressed by a ruled-based fuzzy traversability index to quantify the ease of travel of a terrain by a mobile robot, which is based on real-time measurements of terrain characteristics retrieved from image data [9]. In their algorithm the four key terrain characteristics were the terrain roughness, slope, discontinuity, and hardness. The four terrain characteristics are combined using fuzzy rules to produce a traversability index that quantifies the ease of travel over the terrain. This index too is expressed as a linguistic fuzzy value, which categorizes the risk of traversability as high, medium and low. In this approach, the type of terrain was not identified. Also, since the initial terrain parameters were obtained from vision sensors, changes in illumination may cause incorrect classifications. This system is more likely for obstacle avoidance than terrain evaluation. Terrain parameter identification via wheel terrain interaction analysis has been investigated by Iagnema *et al.* [10]. The problem has been tackled by determining the soil shear strength from two key terrain parameters; cohesion of the soil and internal friction angle. These parameters were estimated online using a simplified form of classical terramechanics equations. Finally using Coulomb's equation, these two parameters were combined to give the shear strength of the soil. This algorithm has limitations related to sensor noise, also not explicitly distinguishing terrain type.

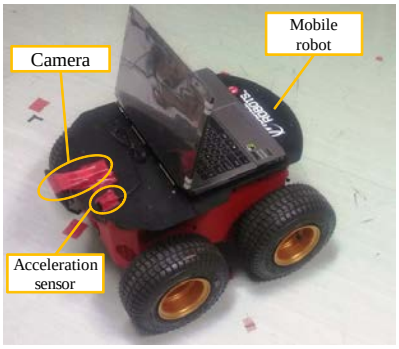

Fig.1 Mobile Robot Pionner 3-AT

In this paper, we propose to demonstrate the possibility of correlation between exteroceptive and proprioceptive information. The objective is to be able to predict from images of the terrain run on, running motion information. The latter will allow to identify terrain type, and thus trigger safety measures if the terrain appears to endanger robot safety and stability. Our test vehicle, a four wheeled mobile robot, shown by **Fig. 1** has an acceleration sensor and a camera mounted on. The vertical acceleration, which is representative of the vibration the robot undergoes is acquired, at the same time the camera captures a still image of the terrain. Features extracted from still images and acceleration signal will be treated as terrain signature, and used to demonstrate a correlation by applying Canonical Correlation Analysis (CCA) framework.

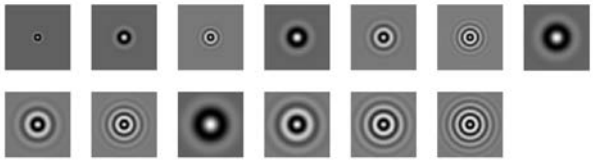

Fig.2 The Shmid filter bank is rotationally invariant and had 13 isotropic, "Gabor-Like" filters

2 Analysis of correlation between motion and image features

2.1 Extraction of image features

Image features are extracted using the Shmid filter bank [11] shown by **Fig. 2**, which consists of 13 rotationally filters described by

$$F(r,\sigma,\tau) = F_0(\sigma,\tau) + \cos\left(\frac{\pi\tau r}{\sigma}\right) e^{-\frac{r^2}{2\sigma^2}}, \tag{1}$$

where τ is the number of cycles of the harmonic function within the Gaussian envelope of the filter, σ is the filter scales, r is the distance between a pixel and the origin of the filter, and $F_0(\sigma,\tau)$ is added to obtain a zero DC component.

First, before applying the filter bank, all images are converted to gray scale, and are normalized to have zero mean and unit standard deviation. This normalization gives partial invariance to linear transformations in the illuminations conditions of the images. Note that, only the region where the robot runs is considered, and thus cropped from the original image as shown by **Fig. 4**, **Fig. 5**, and **Fig. 6**.

Second, all filters of the Shmid filter bank are normalized so that the filter responses lie approximately in the same range. Every filter is convolved with the cropped region, which means that every pixel $I(i,j)$ will have 13 filter responses. Those responses are averaged over all the pixels in the cropped region to form the image feature vector $\boldsymbol{x}$, with $\boldsymbol{x} \in \mathbb{R}^{13}$.

2.2 Extraction of motion features

Let $a_k, t = 1,\cdots,K$ denote the vertical acceleration signal at time step k, where K denotes total time steps for one sequence of a run. Acceleration feature y is calculated by

$$y = \max_{k=1,\cdots,K} a_k - \min_{k=1,\cdots,K} a_k, \tag{2}$$

where a_z is the vertical acceleration signal, and $y \in \mathbb{R}$ is the distance between maximum and minimum amplitudes.

2.3 Correlation analysis

As mentioned above, our goal is to highlight a potential correspondence between image and acceleration features. If any relationship is established, it will be easy to predict running information of the robot from image features. For this CCA is used.

CCA allows to compare two groups of quantitative variables, to determine if they describe a common phenomenon. Let $\boldsymbol{x}_i, i = 1,\cdots,N$ denote image feature vector for the i-th terrain sample, and $y_i, i = 1,\cdots,N$ acceleration feature vector of the same sample, where N represents the number of samples. By re-arranging the data above, matrices are built as follows:

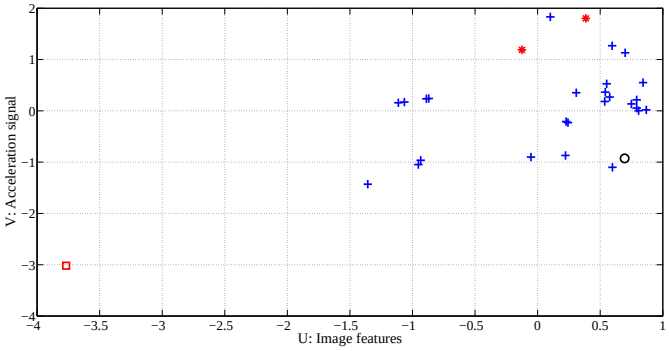

Fig.3 Correspondence between image features and acceleration signal in the (U,V) space

$$\mathbf{X} = [\boldsymbol{x}_1 ... \boldsymbol{x}_N] \in \mathbb{R}^{13\times N}, \tag{3}$$

$$\mathbf{Y} = [y_1 ... y_N] \in \mathbb{R}^{1\times N}. \tag{4}$$

Then, CCA is applied to the above matrices, and coefficients $\mathbf{a} \in \mathbb{R}^{13}$ and $b \in \mathbb{R}$ are computed by maximizing $\rho(\mathbf{a}, b)$, which is expressed as follows:

$$\rho(\mathbf{a},b) = \frac{\mathbf{a}^\top \mathbf{S_{XY}} b}{\sqrt{\mathbf{a}^\top \mathbf{S_{XX}} \mathbf{a}}\sqrt{b \mathbf{S_{YY}} b}}. \tag{5}$$

$\mathbf{S_{XX}} = \frac{1}{N}\tilde{\mathbf{X}}^\top \mathbf{X}$, $\mathbf{S_{YY}} = \frac{1}{N}\tilde{\mathbf{Y}}^\top \mathbf{Y}$, and $\mathbf{S_{XY}} = \frac{1}{N}\tilde{\mathbf{X}}^\top \mathbf{Y}$ are covariance matrices. $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$ are as follows:

$$\tilde{\mathbf{X}} = [\boldsymbol{x}_1 - \bar{\boldsymbol{x}} \ \ldots \ \boldsymbol{x}_N - \bar{\boldsymbol{x}}], \quad \bar{\boldsymbol{x}} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}_i, \tag{6}$$

$$\tilde{\mathbf{Y}} = [y_1 - \bar{y} \ \ldots \ y_N - \bar{y}], \quad \bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i. \tag{7}$$

3 Experiment

3.1 Experimental conditions

Images were gathered with the Logicool Pro C910 with a size of 1944 × 2592 pixels. As mentioned in subsection 2.1, a region with a size of 300 × 400 pixels was cropped from all images. Our acceleration sensor records information with a sampling period equals to 10 ms. Note that for each terrain, 10 acceleration sequences were recorded and averaged before building motion feature vector as described in subsection 2.2. In our case, 30 different environments have been tested. the (σ,τ) pair for generating the filter bank takes values (2,1), (4,1), (4,2), (6,1), (6,2), (6,3), (8,1), (8,2), (8,3), (10,1), (10,2), (10,3), and (10,4).

3.2 Results and discussion

By computing CCA over all data set, a correlation coefficient r equals to 0.579 was obtained. The results are given by **Fig. 3**.

Let us consider the data points resented by a red star. As we can see, a high value for image features *i.e.* texture information of the terrain correspond to a high value for acceleration features representative of the vertical vibration that

Fig.4 Rough natural terrain

Fig.5 Rough natural terrain

the robot endures. This case can be witnessed as the robot traverses rough terrain as shown by **Fig. 4** and **Fig. 5**.

Let us consider the data point represented by a black circle. If we follow the relation established above for this case, even though image feature assume a high value, acceleration features assume a lower value, which means that the vibration that the robot endures is not as important as in the previous case. This situation can be observed when the robot crosses a flat artificial terrain as shown by **Fig. 6**. For some cases as the sample represented by a red square, both image information and acceleration information are not representative of each other. A rich texture information of the terrain does not necessarily translates into an important vibration endured by the robot, and thus considering the terrain as rough would be wrong. It is evident that a more precise terrain discrimination will enable a good terrain classification and thus a good prediction of motion information from image features.

4 Conclusion

In this paper, for predicting mobile robot motion information, we proposed to use the CCA algorithm to establish a correspondence between image feature extracted from images of the scene and vertical acceleration signal. The results show a lack of correlation between both information. This is due to the fact that for some cases, image feature which are richer in information, in this case texture, are not so representative of mobile robot running information, and vise versa. In addition, amount of data for training should be increased.

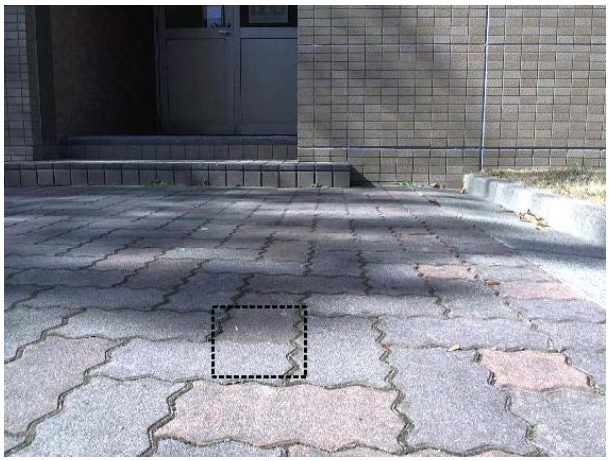

Fig.6 Rigid artificial terrain

Acknowledgement

This research was partly supported by JSPS KAKENHI Grant Number 25330305.

References

[1] D. Sadhukhan, "Autonomous Ground Vehicle Terrain Classification Using Internal Sensors," MS, Master of Science, Department of Mechanical Engineering, Florida State University, 2004.

[2] Manduchi, R. Castano, A. Talukder, A. and Matthies, L., "Obstacle Detection and Terrain Classification for Autonomous Off-Road Navigation," Autonomous Robots, vol. 18, pp. 81-102, 2005.

[3] Dahlkamp, H. Kaehler, A. Stavens, D. Thrun, S. and Bradski, G., "Self-supervised Monocular Road Detection in Desert Terrain," Proceedings of robotics: Science and systems, 2006.

[4] Thrun, S. Montemerlo, M. Dahlkamp, H. Stavens, D. Aron, A. Diebel, J. et al., "Stanley: The robot that won the DARPA Grand Challenge," Journal of Field Robotics, vol. 23, pp. 661-692, 2006

[5] Helmick , D. Angelova, A. and Matthies, L., "Terrain Adaptive Navigation for Planetary Rovers," Journal of Field Robotics, vol. 26, pp. 391-410, 2009.

[6] Goldberg, S. B. Maimone, M. W. and Matthies, L., "Stereo vision and rover navigation software for planetary exploration," IEEE Aerospace Conference Proceedings, vol. 5, pp .2025-2036, 2002.

[7] Inotsume, H. Sutoh, M. K. Nagaoka, Nagatani, K. and Yoshida, K., "Evaluation of the Reconfiguration Effects of Planetary Rovers on their Lateral Traversing of Sandy Slopes," 2012 Ieee International Conference on Robotics and Automation (Icra), pp. 3413-3418, 2012.

[8] Brooks, C. A. Iagnemma, K. D. and Dubowsky, S., "Visual wheel sinkage measurement for planetary rover mobility characterization," Autonomous Robots, vol. 21, pp. 55-64, 2006.

[9] Howard,A. and Seraji, H., "Vision-based terrain characterization and traversability assessment," Journal of Robotic Systems, vol. 18, pp. 577-587, Oct 2001.

[10] Iagnemma, K. Shibly, H. and Dubowsky, S. "On-line terrain parameter estimation for-planetary rovers," IEEE International Conference on Robotics and Automation, 2002, vol. 3, pp. 3142-3147.

[11] Schmid, C. "Constructing models for content-based image retrieval," in Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, 2001, pp. 39-45 vol.2.