A New Terrain Recognition Approach for Predictive Control of Assistive Devices using Depth Vision

Ali H. A. Al-dabbagh and Renaud Ronsse

Abstract— **Vision based systems for terrain detection play important roles in mobile robotics, and recently such systems emerged for locomotion assistance of disabled people. For instance, they can be used as wearable devices to assist blind people or to guide prosthesis or exoskeleton controller to retrieve gait patterns being adapted to the executed task (overground walking, stairs, slopes, etc.). In this paper, we present a computer vision-based algorithm achieving the detection of flat ground, steps, and ramps using a depth camera. Starting from point cloud data collected by the camera, it classifies the environment as a function of extracted features. We further provide a pilot validation in an indoor environment containing a rich set of different types of terrains, even with partial occlusion. The paper further shows that our system needs less computational resources than recently published concurrent approaches, owing to the original transformation method we developed.**

I. INTRODUCTION

The biomechanics of locomotion reveals that leg joints often produce mechanical energy in various locomotion tasks like walking or stair ascending. Moreover, each of these tasks corresponds to different kinetic and kinematic patterns, even in steady-state. As consequence, active prostheses or exoskeletons are necessary to provide disabled users with the capacity to display healthy gait patterns. The control of such devices should be adaptive to the task being executed (walking, taking stairs, etc.) and several researchers recently developed such adaptive algorithms, for instance by using finite state controllers [1]. However, switching from one locomotion mode to another, requires a timely transition, sometimes even before the critical time or event. Interestingly, computer vision-based terrain detection systems can provide such predictive information about the surrounding environment for some steps in advance. More specifically, systems with depth sensing can accurately detect and locate environmental features in 3D spaces.

II. METHODS

We implemented a computer vision-based algorithm for terrain recognition using a depth camera. The raw point cloud data obtained from the depth camera passes through multiple processing steps to extract environmental features, and thus achieve terrain detection. These steps are shown in Fig.1. First data are down-sampled and cropped in order to reduce computational load, since not all points bring a significant part of new information. Then, for each point, a normal vector is estimated from its neighboring points based on a Principal Component Analysis (PCA) [2]. After that, the point cloud data with their normals are passed to a clustering and plane fitting step. The point cloud data is clustered into several groups using a regional growing algorithm. It starts from a point with minimum curvature called *seed*, and expands the region to include neighboring points having quasi-parallel normals and similar curvature [2]. Next, a plane is fitted over each cluster using the RANdom SAmple Consensus (RANSAC) algorithm [3]. At the end of this step, the environment is thus segmented into several planes, whose 3D location is given by their centroid, and orientation is given by the orientation of their averaged normal vector. However, point cloud data acquired from a depth sensor are provided in a coordinate frame being centered on and oriented with the camera. This causes issues for proper in terrain detection, for instance in order to discriminate between a flat and a sloppy ground. We mechanically fixed an IMU to the depth sensor in order to retrieve an estimate of its global orientation in space.

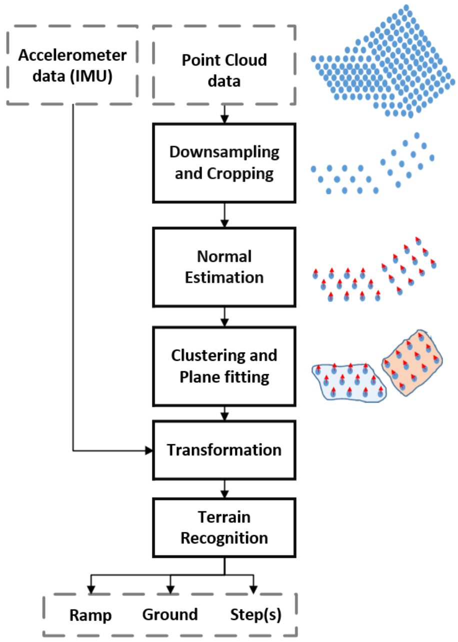


Fig. 1. Block diagram illustrating the overall terrain detection algorithm. Point cloud data acquired from a depth camera and its orientation obtained from an IMU sensor are its inputs. It provides a detection of the locomotion terrain(s) , i.e. flat ground, stairs, and/or ramp, as output

Ali H. A. Al-dabbagh and Renaud Ronsse are with Institute of Mechanics, Materials, and Civil Engineering; Institute of Neuroscience; and Louvain Bionics; Université catholique de Louvain, Louvain-la-Neuve, Belgium (e-mails: ali.aldabbagh@uclouvain.be, renaud.ronsse@uclouvain.be).

This work was supported by the Conseil de l'action internationale of UCLouvain

The feature of the identified planes can then be rotated in the inertial frame. This approach is expected to be less computationally greedy than more classical transformations rotating the whole point cloud before feature extraction [2], [4], [5]. Finally, a classification tree is used for terrain recognition. It tests the transformed features of each plane with respect to several constraints, in order to detect the ground, ramp and step(s). We assigned specific colors to each detected type of terrains for visualization purpose. The algorithm was implemented in the Robot Operating System (ROS), and using the Point Cloud Library (PCL) [6]. The system was tested in a real indoor environment with a Kinect sensor held by hand.

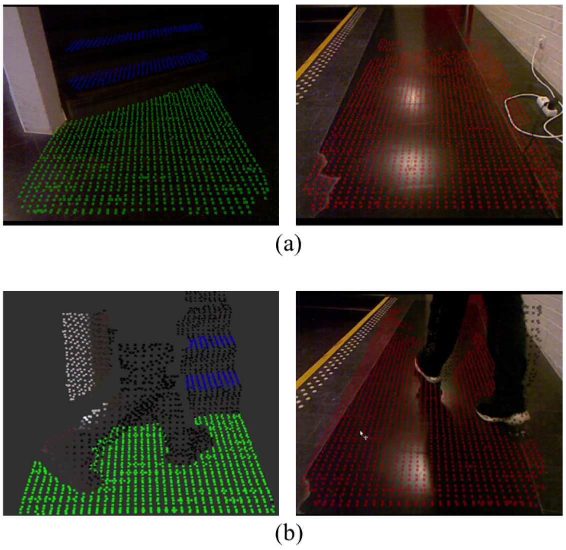
III. RESULTS AND DISCUSSION

Samples of the recognition results are shown in Fig.2. Results displayed in this figure, and many others, show that our methodology successfully detects the different terrains. Moreover, the algorithm proved to be robust even if the terrain is partially occluded by an obstacle. This is important since real urban environment will likely be cluttered with other people or obstacles. More precise quantification of these performance metrics is ongoing.

In a second set of experiments, we quantified the computational cost of our algorithm, with a fixed camera. We further compared it with a more standard approach consisting in rotating the point cloud back to the inertial frame before the feature extraction. This processing time was measured as a function of the cloud size, and different sizes were obtained by successive down-sampling, thus generating low and high density clouds. For each size, ten trials were recorded for both algorithms. Fig.3 reports these trials, and the mean computational time. In sum, our approach with transformation of features only needs less computational time than a more classical one which transforms the whole cloud. In general, our method is about 20% cheaper in computational resources than an "all cloud transformation" method, at least for cloud sizes above 10^5 points.

IV. CONCLUSION

In this paper, a new terrain detection algorithm based on a depth sensing camera is presented. The system detects the main ground, steps and ramps even if they are partially occluded. It requires less computational time because we only transformed the plane features rather than the whole point cloud. Consequently, this approach is promising for wearable systems with limited processing capabilities. In the near future, the system will be tested with different subjects and in several indoor and outdoor environments, in order to precisely quantify its detection abilities.



(a)

(b)

Fig. 2. Terrains recognition samples. Red points for ramp detection, blue points for capture step detection, and green points capture the main ground detection. (a) Real environment. (b) terrains recognition while they are partially occluded.

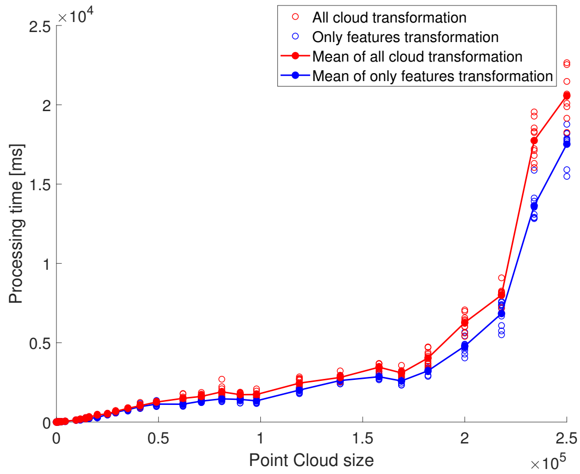


Fig. 3. Processing time for various point cloud sizes. Two different transformation methods were tested: red points denote trials were all points of the cloud are rotated in the world frame before feature extraction, while the blue point correspond to our approach, where we extract features in the camera frame and rotated only these extracted features (plane centroid and orientation). Each dot stands for one recorded run time per frame. The plain lines connect the mean of recorded run times

REFERENCES

[1] M. R. Tucker, J. Olivier, A. Pagel, H. Bleuler, M. Bouri, O. Lambercy, J. del R Millán, R. Riener, H. Vallery, and R. Gassert, "Control strategies for active lower extremity prosthetics and orthotics: a review," *Journal of neuroengineering and rehabilitation*, vol. 12, no. 1, p. 1, 2015.

[2] A. Perez-Yus, D. Gutiérrez-Gómez, G. Lopez-Nicolas, and J. Guerrero, "Stairs detection with odometry-aided traversal from a wearable rgb-d camera," *Computer Vision and Image Understanding*, vol. 154, pp. 192–205, 2017.

[3] R. S. Consensus, "A paradigm for model fitting with applications to image analysis and automated cartography," *MA Fischler, RC Bolles*, vol. 6, pp. 381–395, 1981.

[4] N. E. Krausz, T. Lenzi, and L. J. Hargrove, "Depth sensing for improved control of lower limb prostheses," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 11, pp. 2576–2587, 2015.

[5] K. Zhang, C. Xiong, W. Zhang, H. Liu, D. Lai, Y. Rong, and C. Fu, "Environmental Features Recognition for Lower Limb Prostheses Toward Predictive Walking," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 3, pp. 465–476, 2019.

[6] R. B. Rusu and S. Cousins, "Point cloud library (pcl)," in *2011 IEEE international conference on robotics and automation*, 2011, pp. 1–4.