State of the art of ground plane detection in 3D applications: A systematic review

K. Habib[1] and S. Czarnuch[1,2]
Memorial University
[1]Faculty of Engineering; [2]Faculty of Medicine
St. John's, NL

***Abstract*- Knowledge of the ground plane as a reference plane is critical to developing human tracking systems in 3D. This is particularly true in complicated scenarios where the orientation and location of the sensor, and the number and purpose of a room's occupants are not known by an automated system. Many researchers have investigated methods of detecting the ground plane. However, little is known about how existing methods can be applied to these complicated tracking scenarios. In this study we present results of a literature review of recent studies relevant to ground plane detection in complicated 3D tracking scenarios. RANSAC, Hough transform, and V-disparity have emerged as effective methods of ground plane detection. Furthermore, utilizing moving objects in the scene, vertical or horizontal surfaces have also very recently been proposed as viable methods. Results of this review will inform the development of a novel ground plane detection method for complicated 3D scenarios.**

Index Terms—Ground plane detection, systematic review, computer vision.

I. Introduction

Many researchers have been working on ground plane detection over the last decades for various areas of interest. Automated ground plane detection is an interesting area and exists in many different scopes. For example, automatic robotic movements, roaming of unmanned vehicles, stair detection, human fall detections, human tracking, human gait analysis or whenever a fixed reference plane is necessary, ground plane detection is the primary step. Ground plane detection is commonly used for obstacle detection for robot mobilization ([1–10]) human fall detection ([11, 12]), stair case detection for visually impaired people ([13–15]), and even automated gait analysis with multiple sclerosis ([16]).

Historically, 2D camera based approaches using appearance, homography, or optical flow methods have been well researched. More recently, ground plane detection from depth images obtained using 3D sensors like Microsoft Kinect has become very popular due to the accuracy and robustness of the sensors. Yet, despite the abundance of research into ground plane detection methods, a comprehensive review of both the history and the state of the art remains lacking.

II. Objectives/contributions

The primary objective of this work is to provide a review of 2D and 3D automated ground plane detection methods. Notably, most of the studies reviewed here require an initial assumption to detect the ground plane. For example, the sensor is aligned with the ground plane or fixed, the ground plane is visible, largest on the scene or horizontal and all the obstacles reside on the floor or robots/vehicles/humans move along the ground plane. The outcome of the review is intended to inform development of a system that can detect the ground plane when none of the reviewed assumptions is guaranteed, the tilt angle and location of the camera is unknown, and humans are not previously identified in the scene.

III. Methodology

Relevant publications included in this study were identified by conducting online searches using Google Scholar, SpringerLink, ScienceDirect, Microsoft Academic Research, Scopus and Memorial University Online Library. Some publications were obtained by contacting the authors directly (e.g., by e-mail and message on ResearchGate).

The keywords used (separately and conjointly) to identify publications for this review were:

- Ground plane detection/ "Ground plane detection"
- Ground floor estimation, floor detection
- 3D RGB sensor, Kinect
- Homography, appearance, image, optical flow
- Monocular, stereo, single and wide-angle camera
- RANSAC, Hough transform, V-disparity
- Ground truth detection

In the initial phase of the review the keywords of the papers and publications found in the searches were scanned, and abstracts and conclusions were reviewed. For the second phase a set of inclusion and exclusion criteria were applied to identify the prospective papers found in the first phase. For example, different methodologies that were used, applications, unique and effective methods and significant improvement of existing methods to detect the ground plane. Importance was given to the journals with higher impact factor and conference papers with higher citation index or citations numbers. Finally, publications satisfying all inclusion and exclusion criteria were included in the final phase; the full review.

During the review of an included publication, secondary publications were identified by searching cited references. As well, potentially relevant papers were identified by seeking 'citied by' articles or updated versions of the papers. These secondary publications were processed through the first and second phase of our methodology and considered for inclusion.

IV. Results

Our initial search using the individual keywords returned approximately 100,000 results. Combining multiple keywords and setting the range to the last two decades and an initial review of abstract and conclusion reduced the number to roughly 300 publications during the first phase. Of these publications, 150 were identified in the second phase as relevant to the study based on methodologies, applications, instrument used and improvements. During the final phase, 44 papers were selected for complete review based on relevancy, impact factor, citation index and improvements.

The review identified five main computer vision ground plane detection methods: 1) appearance; 2) homography; 3) motion or optical flow; 4) geometrical and structural; and 5) 3D/RGB-D. We give a detailed review of these approaches.

Appearance based method:

In appearance based methods, visual properties such as colour and texture are collected from an indoor environment to detect the ground floor. Lorigo et al. [1] used colour information (RGB and HSV colour) and intensity gradients (brightness) to detect the ground plane. Ulrich and Nourbakhsh [17] improved Lorigo's model using passive monocular colour vision and then classified each image pixel by determining if the pixel was locating on an obstacle or on the ground. Unlike Lorigo's model, they considered the fact that the ground was not always obstacle free. However, their model required a relatively flat surface, no overhanging obstacles and colour obstacle appeared different in colour to the ground. Colour information also used by Lancer and Veloso [18] to detect obstacles by comparing the colour of obstacles to a ground plane of known colour. Challenge arose when objects and walls were similar in colour to the ground, making it difficult to detect the ground plane. Elleuch et al.[19] extended Ulrich's model by including a texture feature with colour to resolve the issue. Their model successfully detected walls and corridors of the same colour as the ground, but did not include shape features.

Homography based method:

Homography, a method of ground plane detection popularized over the past decade, utilizes a transformation matrix that relates the pixels/points of a plane from two different views (i.e., points that lie in the same plane share a transformation). Commonly, the points near the ground are tracked to get the cues and used to distinguish the floor and obstacles utilizing the floor homography error [2]. To detect the ground plane using homography, researchers have used a number of methods including sparse feature detection to separate the floor and objects [3], dense point features or pixel-to-pixel correspondence by combining appearance and homography cues ([20]), Kanade-Lucas-Tomas feature [21], [22], modified expectation maximization algorithm and SIFT algorithm [4], inverse perspective mapping (IPM) method combined with colour segmentation and Speeded Up Robust Features (SURF) [23] or IPM method to identify pixels on images as belonging to the ground or objects [2]. Cui et al. [21] utilized a planer homography method to detect and segment the floor by identifying plane normals from motion fields in image sequences and using the data to segment the floor.

Motion based or optical flow based method:

An optical flow field is defined as a vector of pixel speeds in an image sequence. In the case of optical flow based ground plane detection, it is usually considered that points lying of the same plane will have a coherent motion pattern which is different than other pixels of the image[22] and the motion can be caused by a moving objects or by the motion of the camera [24]. Optical flow methods have been successfully used to detect the ground plane[21], [24], [25]. Notably, Kumar et al. [22] used a combination of sparse optical flow to determine point correspondence between two successive frames and planar homography to detect the ground plane followed by graph based segmentation for ground floor extraction.

Geometric and structural methods:

Hoang et al. [26] used vertical planes to detect the ground floor, exploiting that walls are perpendicular to the ground plane. Similarly Pears and Liang [10] and Zhou and Li [3] utilized corner information for ground plane detection rather than walls. Although these approaches successfully identified the ground planes, these algorithms faced limitations where the walls were difficult to detect, vertical edges were not visible, out of the range from the camera view or hidden by furniture or other obstacles. Vaz and Ventura [27] tried to overcome these limitations with a ground detection algorithm that was able to be used in semi-structured environments without visible walls, relying on the transition between the ground and other non-planar structures. To detect the ground plane, first they identified the potential ground point cloud in the sensor space, and estimated the ground plane model. Based on the plane parameters, they evaluated the transformation between the previous and current frame coupled with the ground. Following an initialization of the ground plane, the algorithm was capable of estimating the camera orientation and location.

To overcome a planar or horizontal ground dependency, Ess et al.[28] assumed that all the objects resided on the ground plane and the ground plane was not fixed. This assumption allowed consideration of hilly, non-planar ground terrain and tilted camera position. Their proposed method detected the ground plane and pedestrians in a Bayesian network by joint estimation, allowing ground detection in empty scenes by depth measurements and in crowed scenes by large number of constrained object detection. The approach didn't utilize environmental information and was computationally expensive.

Giordano et al. [29] proposed two strategies to determine the ground plane. The first method was based on homography constraints, which depended on correspondences across distant image frames. The second method considered the scene structure, and used depth maps to cluster and extract best fitting planes from 3D point clouds. Comparing methods in an extensive set of simulations suggested that the structure-based method outperformed the homography-based approach, but that their methods need improvements in the presence of multiple planes and use of different kinds of visual features like discrete/dense image moments.

Other researchers have used approaches like LIDAR (McDaniel 2010 [30]) or employed Markov model (Dragon and Van Gool 2014 [31], Kumar et al. 2014 [32]) for ground plane detection. Due to the limitation of time and restriction of page numbers these methods were not included in this paper.

RGB-D sensor or 3D point cloud based approach

3D methods of ground plane detection that utilize RGB-D or 3D point cloud data are the most current and state-of-the-art. Development of these approaches began with simple plane detection but have recently been extended to actual ground plane identification. The most popular 3D methods of plane detection are RANdom SAmples Consensus (RANSAC), Hough transform, and V-disparity approaches. RANSAC is one of the earliest plane detection algorithm, proposed by Fischler & Bolles [33] and is used directly and as a part of many complex ground plane detection methods. However, this method is simply useful for detecting planes in 3D data, and cannot disambiguate the ground plane from any other plane. Furthermore, the iterative nature of the algorithm can at times be computationally costly. Similarly, a Hough transformation approach was proposed by Borrmann et al. [34] to detect planes in 3D point clouds, but the computational cost of the method was too high to be practical. To improve the plane detection methods, Yoo et al. [35] suggested a fast system of multiple plane detection using depth map data by computing local normal vectors of points and classifying 3D point cloud data. Instead of iterating through each point cloud (like RANSAC), the cloud data were sampled uniformly on the depth map data and each point was checked to confirm if it was in the same plane. These plane detection algorithms have become the basic tools used to subsequently segment the ground plane from the scene.

To detect the ground plane directly, Jin et al. [36] proposed an algorithm that identified the flattest/most planner surface on the depth map by finding valid seed patches over the whole depth map. Instead of generating random points like RANSAC, their approach used dynamic threshold and surface function to expand the point growing process until no new point fit into the planar surface. Rodríguez et al. [37] used the RANSAC method with some filtering techniques to estimate the ground plane by analyzing 3D point cloud data. They initially assumed that the camera was fixed and the ground plane was the largest area in the bottom part of the images. Their proposed method defined a polar cumulative grid filled with 3D points on the ground plane and then filtered out any pixels located above a certain height as possible obstacles. The remaining points were used for plane parameter estimation using RANSAC. This technique proved to be robust, but the experiments showed errors in cases of glossy ground surfaces and in some cases detected the wrong ground plane.

Tang et al. [13] proposed a plane-based approach to detect staircases to assist visually impaired people. They detected staircases by modeling individual steps as horizontal planes in 3D space and identified them by a modified RANSAC method (proposed by Nister [38]). As a first step of their algorithm, they located the ground plane in the scene by a rough estimation of the height of the sensor from the ground and the orientation of the gravity vector from an imbedded sensor inclinometer. Points estimated to be part of the ground plane were used with the RANSAC method for hypothesis generation. Instead of using a large number of points in a larger plane, they used N (generally 3) number of points for the robust detection of planes to reduce the computational cost. If the plane hypothesis was not parallel to the ground plane, the data was discarded. For the remaining hypothesises, voting was performed using random subsets of points to identify the best performing hypothesis until the last hypothesis remained, which was accepted as the final plane model. Their proposed method allowed filtering noisy and poor fitting points and planes before being used in the plane hypothesis generation and voting stage, thus improving the cost of the algorithm. Unfortunately, their model depended on an estimate of the height of the sensor to the floor and that the horizontal plane was always visible. Vlaminck et al. [14] and Perez-Yus et al. [15] used the dataset of Tang et al. [13] to improve ground plane modeling and overcome uncertainties. Perez-Yus et al. [15] used RANSAC to find the largest plane based on the approximate location information from the chest-mounted camera. They found all planes, and analyzed each plane to determine the relative distance and orientation of each plane with respect to the camera until a valid ground plane was found. They assumed that the orientation of the normal with the camera and the distance of the plane to the camera were within valid ranges and the floor points were closer than a certain threshold. Similarly, Vlaminck et al. [14] detected ground plane as a preprocess step for stair detection. Using point cloud data obtained from depth images, they calculated the surface normal for each point leading to a single normal map. The map was computed using the 2D projection of the point cloud by considering two vectors tangential to the local surface at a certain point. These vectors were computed using the neighbouring points and then the normal for that point was computed by taking their cross product. Finally, they transformed the point cloud in the reference system in order to align the ground plane with the xz-plane based on the fact that the ground plane has a normal perpendicular to the xz-plane. The dataset was complemented by data from an accelerometer that was used to determine the orientation of the camera. Using the accelerometer data allowed explicit identification of the camera pose for each frame, allowing the detection of a plane parallel to the ground floor even when the ground was not visible by the camera. Notably, each of these 3D ground plane detection methods require knowledge of the orientation and/or an estimate of the location of the sensor. When the orientation or tilt angle of the camera/sensor is unknown, extra steps need to be taken in order to detect the ground plane. Lang et al. [39] proposed an algorithm to estimate the ground plane with a tilted sensor using a Kernel Density Estimator method. They used probability theory to first estimate the tilt angle of the camera and used this angle to improve the accuracy of the subsequent ground plane detection.

Uncertainty in the location of the sensor further complicates ground plane detection. Kepski and Kwolek [11], assuming a person could first be identified in the scene, focused on human fall detection by calculating the distance of the person to the ground plane. They established a method that automatically extracted the ground plane in depth images using V-disparity method, Hough transform and the RANSAC algorithm. The original work of V-disparity, proposed by Labayrade et al. [40], used disparity maps between two stereo images. Following their method, Kepski and Kwolek [11] calculated the disparity image from the depth map and used the Hough transformation to extract the line corresponding to the floor pixels. Based on the known height of the sensor from the floor, tilt angle of the sensor and the extracted line, the pixels as a part of the floor were identified and transformed to 3D point cloud data. Finally, the ground plane was identified using RANSAC. With the assumption that the floor was the largest part of the scene, Rougier et al. [12] developed a similar approach using V-disparity images and Hough transformation for ground plane detection, but didn't use the RANSAC method because of the computational cost. After the floor pixels were detected from the V-disparity image and the line corresponding to the ground plane was detected from Hough transformation, these pixels and their known depths were used to recover the 3D plane equation and compute the parameters of the ground plane using a least squares fit of the 3D detected points. In contrast to their method, rather than considering the ground plane as the largest horizontal plane, Kircali and Tek [41] modelled ground planes with some inclination (or declination) by determining the degree of the curve of the plane based on an exponential curve fit. They used their approach when the viewing angle was fixed and also in the dynamic scenario where the sensor viewing angle changed in every frame. They compared their results with the V-disparity approach which relied on the linear increase of disparity values and the fitting of a ground plane model line, and showed that their method was a useful approach in cases of difficult scenarios and dynamic environments. However, their approach largely depends on the curve fitting procedures and can produce errors if the fitting is unsuccessful. Emaduddin et al. [42] developed a recursive RANSAC segmentation algorithm that divided the initial point cloud into multiple regions of interest and estimated the dominant and sub-dominant plane models within a ground plane. They evaluated their approach on 3D data from multiple sensors, showing an improved computational time and selection of floor sections. Ground plane detection has also been used as a primary step for human tracking. For example, with the assumption that people walk on the ground plane, Munaro et al. [43] proposed an algorithm that detected the ground plane from RGB-D data. They calculated the plane coefficients with RANSAC-based least square method and eliminated all the inliers within a threshold distance. Updating the ground plane equation at every frame allowed them to detect real time changes within the floor, like floor slope and camera oscillation. Similarly, as a basis of human tracking, Jafari et al. [44] computed the occupancy map from point cloud data based on the camera height and by excluding the high density points. The ground plane was estimated from the remaining points by plane fitting using RANSAC. However, both approaches required a priori knowledge of the position and orientation of the camera.

Czarnuch and Ploughman [16] proposed a method to detect the ground plane in 3D point clouds without knowledge of either the sensor location and orientation, or previous segmentation of humans. Their method required that a human or object move in a motion predominantly parallel to the ground plane. Euclidean clustering was used in each frame of captured data, and the trajectories of clusters that moved across successive frames were computed. RANSAC was employed to identify all planes in the scene. Any planes that were parallel to the trajectories of the moving objects were then identified as potential ground planes. The ground plane was explicitly detected if the moving objects were all above or in contact with a potential plane. However, their approach was only evaluated on a small data sample in ideal environments.

V. Discussion/Conclusion

This paper presents a systematic review of all modern 2D and 3D methods of automated ground plane detection. Specifically, we seek to use this review to inform a solution to the problem of ground plane detection in a complicated condition were three key pieces of information are unknown: 1) the location of the sensor; 2) the orientation of the sensor; and 3) what objects in the scene are humans. All of the approaches reviewed except [16] require at least one of the above pieces of information are known to be successful. Only the preliminary work of Czarnuch and Ploughman [16] shows promise of implicitly or explicitly identify the ground plane without any of the above information. Notwithstanding the limitations of the reviewed methods, recent developments in 3D methods show the most promise at identifying the ground plane in this complicated condition because of their robustness at both plane detection and their inherent 3D geometric data.

Acknowledgment

Authors would like to thank the NSERC Discovery Grant program for funding this work.

References

[1] L. M. Lorigo, R. A. Brooks, and W. E. L. Grimsou, “Visually-guided obstacle avoidance in unstructured environments,” in *Intelligent Robots and Systems, 1997. IROS’97., Proceedings of the 1997 IEEE/RSJ International Conference on*, 1997, vol. 1, pp. 373–379.

[2] T.-J. Lee, D.-H. Yi, D.-I. Cho, and others, “A Monocular Vision Sensor-Based Obstacle Detection Algorithm for Autonomous Robots,” *Sensors*, vol. 16, no. 3, p. 311, 2016.

[3] J. Zhou and B. Li, “Robust ground plane detection with normalized homography in monocular sequences from a robot platform,” in *2006 International Conference on Image Processing*, 2006, pp. 3017–3020.

[4] D. Conrad and G. N. DeSouza, "Homography-based ground plane detection for mobile robot navigation using a modified em algorithm," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, 2010, pp. 910–915.

[5] Y. Kim and H. Kim, "Layered ground floor detection for vision-based mobile robot navigation," in *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on*, 2004, vol. 1, pp. 13–18.

[6] Y. Li and S. T. Birchfield, "Image-based segmentation of indoor corridor floors for a mobile robot," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, 2010, pp. 837–843.

[7] M. Wang, Y. Chung, and H. Lin, "A self-localization technique for mobile robots using image-based ground plane detection," in *Proceedings of the 2009 International Conference on Service and Interactive Robotics (SIRCon)*, 2009.

[8] J. Cunha, E. Pedrosa, C. Cruz, A. J. Neves, and N. Lau, "Using a depth camera for indoor robot localization and navigation," *DETI/IEETA-University of Aveiro, Portugal*, 2011.

[9] J. Biswas and M. Veloso, "Depth camera based indoor mobile robot localization and navigation," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, 2012, pp. 1697–1702.

[10] N. Pears and B. Liang, "Ground plane segmentation for mobile robot visual navigation," in *Intelligent Robots and Systems, 2001. Proceedings. 2001 IEEE/RSJ International Conference on*, 2001, vol. 3, pp. 1513–1518.

[11] M. Kepski and B. Kwolek, "Human Fall Detection Using Kinect Sensor," in *Proceedings of the 8th International Conference on Computer Recognition Systems CORES 2013*, vol. 226, R. Burduk, K. Jackowski, M. Kurzynski, M. Wozniak, and A. Zolnierek, Eds. Heidelberg: Springer International Publishing, 2013, pp. 743–752.

[12] C. Rougier, E. Anvient, J. Rousseau, M. Mignotte, and J. Meunier, "Fall detection from depth map video sequences," Intl. Conf. on Smart Homes and Health Telematics, pp. 121-128, 2011.

[13] T. J. J. Tang, W. L. D. Lui, and W. H. Li, "Plane-based detection of staircases using inverse depth," in *Australasian Conference on Robotics and Automation (ACRA)*, 2012.

[14] M. Vlaminck, Q. L. Hiep, H. Vu, P. Veelaert, W. Philips, and others, "Indoor assistance for visually impaired people using a RGB-D camera," in *2016 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, 2016, pp. 161–164.

[15] A. Perez-Yus, D. Gutierrez-Gomez, G. Lopez-Nicolas, and J. J. Guerrero, "Stairs detection with odometry-aided traversal from a wearable RGB-D camera," *Computer Vision and Image Understanding*, May 2016.

[16] S. Czarnuch and M. Ploughman, "Automated gait analysis in people with multiple sclerosis using two unreferenced depth imaging sensors: Preliminary steps," in *Proceedings of the 29th International Conference on Image and Vision Computing New Zealand, IVCNZ*, 2014, pp. 19–21.

[17] I. Ulrich and I. Nourbakhsh, "Appearance-based obstacle detection with monocular color vision," in *AAAI/IAAI*, 2000, pp. 866–871.

[18] S. Lenser and M. Veloso, "Visual sonar: Fast obstacle avoidance using monocular vision," in *Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on*, 2003, vol. 1, pp. 886–891.

[19] J. F. Elleuch, I. K. Kallel, and D. S. Masmoudi, "New Ground Plane Segmentation Method for Electronic Cane," *Journal of Image and Graphics*, vol. 1, no. 2, pp. 72–75, 2013.

[20] E. Fazl-Ersi and J. K. Tsotsos, "Region classification for robust floor detection in indoor environments," in *International Conference Image Analysis and Recognition*, 2009, pp. 717–726.

[21] X.-N. Cui, Y.-G. Kim, and H. Kim, "Floor segmentation by computing plane normals from image motion fields for visual navigation," *International Journal of Control, Automation and Systems*, vol. 7, no. 5, pp. 788–798, Oct. 2009.

[22] S. Kumar, A. Dewan, and K. M. Krishna, "A bayes filter based adaptive floor segmentation with homography and appearance cues," in *Proceedings of the Eighth Indian Conference on Computer Vision, Graphics and Image Processing*, 2012, p. 54.

[23] C.-H. Lin, S.-Y. Jiang, Y.-J. Pu, and K.-T. Song, "Robust ground plane detection for obstacle avoidance of mobile robots using a monocular camera," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, 2010, pp. 3706–3711.

[24] C. Braillon, C. Pradalier, J. L. Crowley, and C. Laugier, "Real-time moving obstacle detection using optical flow models," in *2006 IEEE Intelligent Vehicles Symposium*, 2006, pp. 466–471.

[25] A. Talukder and L. Matthies, "Real-time detection of moving objects from moving vehicles using dense stereo and optical flow," in *Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*, 2004, vol. 4, pp. 3718–3725.

[26] V.-N. Hoang, T.-H. Nguyen, T.-L. Le, T.-H. Tran, T.-P. Vuong, and N. Vuillerme, "Obstacle detection and warning system for visually impaired people based on electrode matrix and mobile Kinect," *Vietnam Journal of Computer Science*, Jul. 2016.

[27] M. Vaz and R. Ventura, "Real-time ground-plane based mobile localization using depth camera in real scenarios," *Journal of Intelligent & Robotic Systems*, vol. 80, no. 3–4, pp. 525–536, 2015.

[28] A. Ess, K. Schindler, B. Leibe, and L. Van Gool, "Object detection and tracking for autonomous navigation in dynamic environments," The International Journal of Robotics Research, vol. 29, pp. 1707–1725, 2010.

[29] P. R. Giordano, R. Spica, and F. Chaumette, "An active strategy for plane detection and estimation with a monocular camera," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, 2014, pp. 4755–4761.

[30] M. W. McDaniel, T. Nishihata, C. A. Brooks, and K. Iagnemma, "Ground plane identification using LIDAR in forested environments," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, 2010, pp. 3831–3836.

[31] R. Dragon and L. Van Gool, "Ground plane estimation using a hidden markov model," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 4026–4033.

[32] S. Kumar, M. S. Karthik, and K. M. Krishna, "Markov Random Field based small obstacle discovery over images," in 2014 IEEE International Conference on Robotics and Automation (ICRA), 2014.

[33] 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.

[34] D. Borrmann, J. Elseberg, K. Lingemann, and A. Nüchter, "The 3D Hough Transform for plane detection in point clouds: A review and a new accumulator design," *3D Research*, vol. 2, no. 2, Jun. 2011.

[35] H. W. Yoo, W. H. Kim, J. W. Park, W. H. Lee, and M. J. Chung, "Real-time plane detection based on depth map from kinect," in *Robotics (ISR), 2013 44th International Symposium on*, 2013, pp. 1–4.

[36] Z. Jin, T. Tillo, and F. Cheng, "Depth-map driven planar surfaces detection," in *Visual Communications and Image Processing Conference, 2014 IEEE*, 2014, pp. 514–517.

[37] A. Rodríguez, J. J. Yebes, P. Alcantarilla, L. Bergasa, J. Almazán, and A. Cela, "Assisting the Visually Impaired: Obstacle Detection and Warning System by Acoustic Feedback," *Sensors*, vol. 12, no. 12, pp. 17476–17496, Dec. 2012.

[38] D. Nistér, "Preemptive RANSAC for live structure and motion estimation," Machine Vision and Applications, vol. 16, pp. 321–329, 2005.

[39] Y. Lang, H. Wu, T. Amano, and Q. Chen, "An iterative convergence algorithm for single/multi ground plane detection and angle estimation with RGB-D camera," in *Image Processing (ICIP), 2015 IEEE International Conference on*, 2015, pp. 2895–2899.

[40] R. Labayrade, D. Aubert, and J.-P. Tarel, "Real time obstacle detection in stereovision on non flat road geometry through 'v-disparity' representation," 2003, vol. 2, pp. 646–651.

[41] D. Kırcalı and F. B. Tek, "Ground plane detection using an rgb-d sensor," in *Information Sciences and Systems 2014*, Springer, 2014, pp. 69–77.

[42] M. Emaduddin, K. AlMutib, M. AlSulaiman, R. Hedjar, and E. Mattar, "Accurate floor detection and segmentation for indoor navigation using RGB+ D and stereo cameras," in *Proceedings of the 2012 International Conference on Image Processing, Computer Vision, & Pattern Recognition*, 2012, pp. 293–299.

[43] M. Munaro, F. Basso, and E. Menegatti, "Tracking people within groups with RGB-D data," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 2101–2107.

[44] O. H. Jafari, D. Mitzel, and B. Leibe, "Real-time RGB-D based people detection and tracking for mobile robots and head-worn cameras," 2014, pp. 5636–5643.