Stereo-based Terrain Traversability Analysis for Robot Navigation

Geert De Cubber[1], Daniela Doroftei[1],
Lazaros Nalpantidis[2], Georgios Ch. Sirakoulis[3], Antonios Gasteratos[2]

[1]Royal Military Academy
Department of Mechanical Engineering (MSTA)
Av. de la Renaissance 30, 1000 Brussels
Geert.De.Cubber@rma.ac.be, Daniela.Doroftei@rma.ac.be
[2]Democritus University of Thrace
Department of Production and Management Engineering
University Campus, Kimmeria, 671 00 Xanthi, Greece
lanalpa@pme.duth.gr, agaster@pme.duth.gr
[3]Democritus University of Thrace
Department of Electrical and Computer Engineering
University Campus, Kimmeria, 671 00 Xanthi, Greece
gsirak@ee.duth.gr

Keywords: Stereo Vision, Terrain Traversability, Outdoor Mobile Robot Navigation

1. Introduction

Outdoor mobile robots, which have to navigate autonomously in a totally unstructured environment need to auto-determine the suitability of the terrain around them for traversal. Traversability estimation is a challenging problem, as the traversability is a complex function of both the terrain characteristics, such as slopes, vegetation, rocks, etc and the robot mobility characteristics, i.e. locomotion method, wheel properties, etc.

In this paper, we present an approach where a classification of the terrain in the classes “traversable” and “obstacle” is performed using only stereo vision as input data. In a first step, high-quality stereo disparity maps are calculated by a fast and robust algorithm. This stereo algorithm is explained in section 3 of this paper.

Using this stereo depth information, the terrain classification is performed, based upon the analysis of the so-called "v-disparity" image which provides a representation of the geometric content of the scene. Using this method, it is possible to detect non-traversable terrain items (obstacles) even in the case of partial occlusion and without any explicit extraction of coherent structures or any a priori knowledge of

Keywords: Stereo Vision, Terrain Traversability, Outdoor Mobile Robot Navigation
Funding: View-Finder FP6 IST 045541

the environment. The sole algorithm parameter is a single factor which takes into account the robot mobility characteristics. This terrain traversability estimation algorithm is explained in section 4.

The stereo disparity mapping and terrain traversability estimation processes are integrated in an autonomous robot control architecture, proving that the algorithms allow real-time robot control. The results of experiments with this robot navigating on rough outdoor terrain are presented in section 5.

2. Previous Work

2.1.Stereo

Autonomous robots' behavior greatly depends on the accuracy of their decision making algorithms. In the case of stereo vision-based navigation, the accuracy and the refresh rate of the computed depth images, i.e. disparity maps, are the cornerstone of its success [19]. In order to address the demand for real-time operation, robotic applications usually utilize algorithms of low computational payload. Thus, local methods are preferred to the global ones. Dense local methods calculate depth for almost every pixel of the scenery, talking into consideration only a small neighborhood of pixels each time [15].

Muhlmann et al in [11] describe a method that uses the sum of absolute differences (SAD) correlation measure for RGB color images. Applying a left to right consistency check, the uniqueness constraint and a median filter, it can achieve 20 fps for 160x120 pixel images. Another fast SAD based algorithm is presented in [2]. It is based on the uniqueness constraint and rejects previous matches as soon as better ones are detected. It achieves 39.59 fps speed for 320x240 pixel images with 16 disparity levels and the root mean square error for the standard Tsukuba pair is 5.77 .The algorithm reported in [23] achieves almost real-time performance. It is once more based on SAD but the correlation window size is adaptively chosen for each region of the picture. Apart from that, a left to right consistency check and a median filter are utilized. The algorithm is able to compute 7 fps for 320x240 pixel images with 32 disparity levels.

Another possibility in order to obtain accurate results in real-time is to utilize programmable graphic processing units (GPU). In [24] a hierarchical disparity estimation algorithm is presented. This method can process either rectified or non-calibrated image pairs using a local SAD based bidirectional process. This algorithm is implemented on an ATI Radeon 9700 Pro GPU and can achieve up to 50 fps for 256x256 pixel input images.

On the other hand, an interesting but very computationally demanding local method is presented in [22]. It uses varying weights for the pixels in a given support window, based on their color similarity and geometric proximity. However, the execution speed of the algorithm is far from being real-time. The running time for the Tsukuba image pair with a 35x35 pixels support window is about one minute. The error ratio is only 1.29%, 0.97%, 0.99%, and 1.13% for the Tsukuba, Sawtooth, Venus and Map image sets accordingly.

A detailed taxonomy and presentation of dense stereo correspondence algorithms can be found in [15]. Additionally, the recent advances in field as well as the aspect of hardware implementable stereo algorithms are covered in [13].

2.2.Traversability Analysis

Terrain traversability analysis is a research topic which has been in the focus of the mobile robotics community in the past decade, inspired by the development of autonomous planetary rovers and, more recently, the DARPA Grand Challenge. However, already in 1994, Langer et al. [10] computed elevation statistics of the terrain (height difference and slope) and classified terrain cells as traversable or untraversable by comparing these elevation statistics with threshold values. Most of the terrain traversability analysis algorithms employ such a cell-based traversability map, which can be thought of as a 2.5D occupancy grid. The problem with Langer's method was that the traversability was only expressed in binary forms and soon other researchers [17][5] presented solutions to lift this limitation. In [16], Seraji proposed a fuzzy-logic traversability measure, called the Traversability index, which represents the degree of ease with which the regional terrain could be navigated. This degree was calculated on the basis of the terrain roughness, the slope and the discontinuity, as measured by a stereo vision system.

Schäfer et al. presented in [14] a similar stereo-discontinuities based approach without explicit calculation of a traversability map. Other researchers [18][8][6] have embedded the stereo-based terrain traversability analysis in an on-line learning approach. The results of these methods depend greatly on the quality of the training set.

In [21], Ulrich and Nourbakhsch presented a solution for appearance-based obstacle detection using a single color camera. Their approach makes the assumption that the ground is flat and that the region in front of the robot is ground. In [7], Kim et al. present another single-camera traversability estimation method based upon self-supervised learning of superpixel regions.

Besides monocular and stereo vision, laser range finders are a useful sensor for terrain traversability estimation. In [1], Andersen et al. present a method for terrain classification using single 2D scans. More recently, in [20], the Stanford Racing Team utilized a Traversability Map based on data from six laser scanners registered with pose from an unscented Kalman Filter to classify grids as undrivable, drivable, or unknown. Unfortunately, pose error often led to a large error in the 3D data. To correct for this a Markov model was used to probabilistically test for the presence of an obstacle leading to an improved Traversability Map. In addition, parameters of the Markov model where tuned using a discriminative learning algorithm and data labeled through human driving. Data representing where the vehicle traveled was labeled as drivable while areas to the left and right of the vehicle were labeled as non-drivable. This significantly reduced the instances of false positives in the map. Finally, a mixture of Gaussians from RGB vision data was maintained for the drivable area of the Traversability Map. These Gaussians were used by an online learning algorithm to label data beyond the range of the laser map. Stanford's extension of the Traversability map represents perhaps the most sophisticated work in the area to date. However, it should be noted that the problem was formulated as a road following problem and has not been tested in off-road navigation scenarios.

3. Stereo disparity mapping

Stereo disparity is computed using a three-stage local stereo correspondence algorithm. The algorithm utilized is a modified version of the algorithm presented in [12]. It combines low computational complexity with sophisticated data processing. Consequently, it is able to produce dense disparity maps of good quality in frame rates suitable for robotic applications. The structural elements of this algorithm are presented in Figure 1. The main attribute that differentiates this algorithm from the majority of the other ones is that the matching cost aggregation step consists of two sophisticated sub-steps rather than one simple summation. In addition, the disparity selection process is a non-iterative one and the final refinement step is absent.

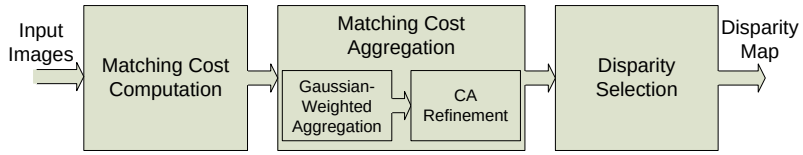

Figure 1: Block diagram of the proposed stereo correspondence algorithm.

The results refinement step is moved inside the aggregation, rather than being an additional final procedure. Instead of refining the results that were chosen through a strict selection process, the proposed algorithm performs a refinement procedure to all the available data. Such a procedure enhances the quality of the results. Thus, the disparity selection step can remain a simple winner-takes-all (WTA) choice. The absence of an iteratively updated selection process significantly reduces the computational payload of this step.

The matching cost function utilized is the truncated absolute differences (AD). AD is inherently the simplest metric of all, involving only summations and finding absolute values. The AD are truncated if they excess the 4% of the maximum intensity value. Truncation suppresses the influence of noise in the final result. This is very important for stereo algorithms that are intended to be applied to outdoors scenes. Outdoors pairs usually suffer from noise induced by a variety of reasons, e.g. lighting differences and reflections. Disparity space image (DSI) is a 3D matrix containing the computed matching costs for every pixel and for all its potential disparity values. The DSI values for constant disparity value are aggregated inside fix-sized square windows. The dimensions of the chosen aggregation window play an important role in the quality of the final result. Generally, small dimensions preserve details but suffer from noise. On the contrast, large dimensions may not preserve fine details but significantly suppress the noise. This behavior is highly appreciated in outdoors robotics applications where noise is a major factor, as already discussed. The aggregation window's dimensions used in the proposed algorithm are 11x11 pixels. This choice is a compromise between real-time execution speed and noise cancelation. The AD aggregation step of the proposed algorithm is a weighted summation. Each pixel is assigned a weight depending on its Euclidean distance from the central pixel. A 2D Gaussian function determines the weight's value for each pixel. The center of the function coincides with the central pixel. The standard deviation is equal to the one third of the distance from the central pixel to the nearest window-border. The applied weighting function can be calculated once and then be applied to all the aggregation windows without any further change. Thus, the computational load of this procedure is kept within

reasonable limits. The DSI values after the aggregation are furthered refined by applying 3D cellular automata (CA). Two CA transition rules are applied to the DSI. The values of parameters used by them were determined after extensive testing to perform best. The first rule attempts to resolve disparity ambiguities. It checks for excessive consistency of results along the disparity axis and, if necessary, corrects on the perpendicular plane. The second rule is used in order to smoothen the results and at the same time to preserve the details on constant-disparity planes. The two rules are applied once. Their outcome comprises the enhanced DSI that will be used in order the optimum disparity map to be chosen by a simple, non-iterative WTA step.

4. Terrain Traversability Estimation

Detecting obstacles from stereo vision images may seem simple, as the stereo vision system can provide rich depth information. However, from the depth image, it is not evident to distinguish the traversable from the non-traversable terrain, especially in outdoor conditions, where the terrain roughness and the robot mobility parameters must be taken into account. Our approach is based on the construction and subsequent processing of the *v-disparity* image [9], which provides a robust representation of the geometric content of road scenes. The v-disparity image is constructed by calculating a horizontal histogram of the disparity stereo image.

Consider 2 stereo frames, as shown in Figure 2*a* and *b*, and the computed disparity image I_D, as shown in Figure 2*c*. Then, the v-disparity image I_V can be constructed by accumulating the points with the same disparity that occur on a horizontal line in the image. Figure 2*d* displays the v-disparity image I_V for the given input images.

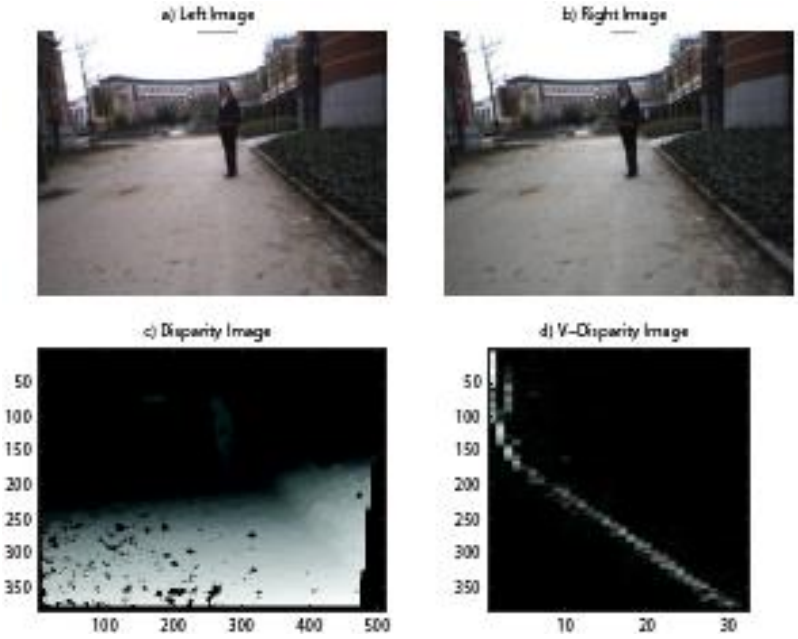

Figure 2: Stereo, Disparity and V-Disparity Images

The classification of the terrain in traversable and non-traversable areas goes out from the assumption that the majority of the image pixels are related to traversable terrain of the ground plane. The projection of this ground plane in the v-disparity image is a straight line, from the top left to the bottom

right of the v-disparity image. Any deviations from this projection of the ground plane are likely obstacles or other non-traversable terrain items.

As such, the processing of the v-disparity image comes down to estimating the equation of the line segment in the v-disparity image, corresponding to the ground plane. This is done by performing a Hough transform on the v-disparity image and searching for the longest line segment. Then, one must choose a single parameter which accounts for the maximum terrain roughness. As this parameter depends only on the robot characteristics, it only needs to be set once. This parameter sets the maximum offset in v-disparity space to be considered part of the ground plane. Any outliers are regarded as obstacles, which enables to compile an obstacle image I_O.

5. Results

5.1. System architecture

As mentioned before, the stereo disparity mapping and terrain traversability estimation processes are integrated in a behavior-based autonomous robot control architecture [4]. This architecture is roughly sketched on Figure 3. As can be noted from Figure 3, the stereo framegrabber and subsequent traversability estimation module form the input for a behavior to direct the robot away from obstacles. This is of course not the only navigational input. The robot uses other on-board sensing equipment such as a differential GPS, an orientation sensor, wheel encoders and visual streams for Visual SLAM, a sonar array for obstacle avoidance, a chemical sensor for detection of chemical contaminants, a heat sensor for avoiding and mapping hot zones and a vision-based human victim detector [3]. All these sensing and processing modalities are combined together in a behavior-based context [4], and a globally optimal robot control command is send to the robot.

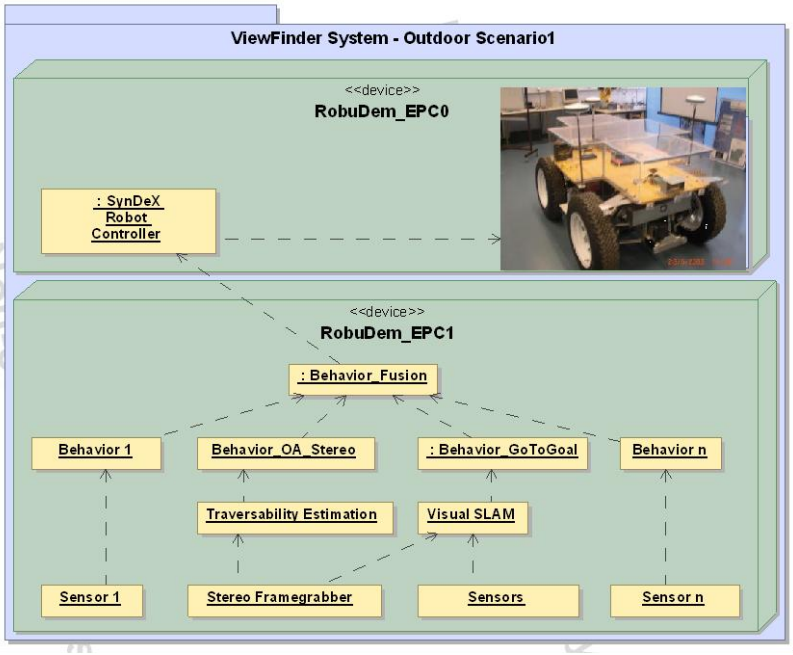

Figure 3: System Architecture

5.2.Stereo results

The stereo algorithm takes as input two images of the scene, captured by a stereo camera. The outcome of the algorithm is a disparity map. A disparity map is usually a grayscale image. The brighter a pixel in the disparity map, the more close the corresponding pixel in the reference image. Consequently, an obstacle could be identified as a group of pixels, which are generally brighter than their neighborhood. The results of the stereo algorithm for various outdoors scenes as well as the corresponding calculated disparity maps are presented in Figure 4.

Black areas, corresponding to regions of the depicted path, occur because the algorithm fails to match correctly such poorly textured areas. In the second and third scenes where there are obstacles, the stereo algorithm successfully highlights them. It can be seen that the used stereo algorithm significantly suppresses the noise, preserving at the same time the crucial details of the scene.

(a) (b) (c)

Figure 4: (a) Left, (b) Right Images of the Stereo Pair and (c) the Computed Disparity Map.

5.3.Traversability Analysis results

Figure 5 presents an example result of the proposed terrain traversability analysis algorithm. The left image shows the V-Disparity image after Hough transform. The red line indicates the largest line segment, corresponding to the ground plane. The two pink lines indicate the region in v-disparity space where pixels are considered part of a traversable region. Terrain corresponding to pixels in v-disparity space in between the two pink lines is considered traversable, otherwise it is considered as an obstacle.

The result of this operation can be judged from the right image of Figure 5, showing the obstacle image. This is a version of the color input image, where false color data corresponding to the disparity is superposed for pixels classified as belonging to non-traversable terrain.

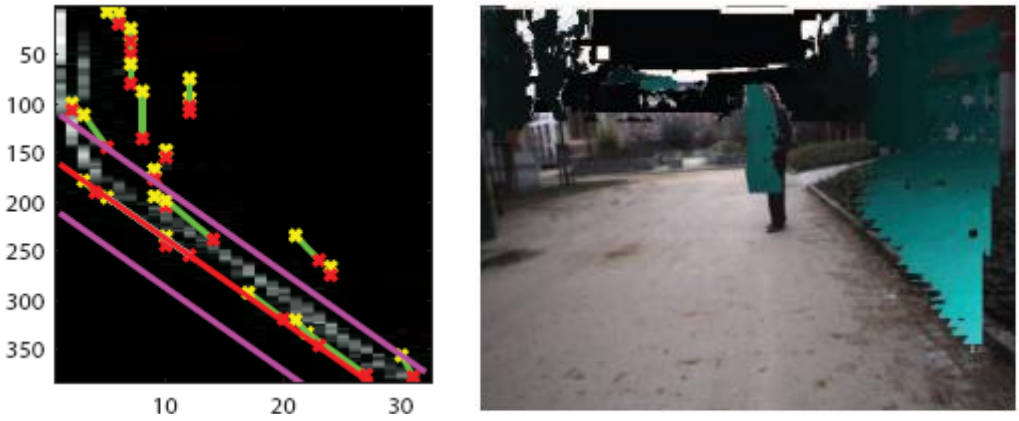

Figure 5: Left: Line-based Analysis of the V-Disparity Image and Right: Resulting Obstacle Image

It may be noted that the lower part of the legs of the person standing in front of the robot were not detected as obstacles. This is due to the choice of the threshold parameter for the ground plane, discussed above. After tests in multiple environments, we used a threshold parameter of 50, which offers a good compromise between a good detection rate and low false positive detection rate.

6. Conclusions

In this article, we have presented a stereo vision algorithm and a stereo-based terrain traversability estimation algorithm. Combined, these approaches make it possible to robustly classify the terrain of outdoors scenes in traversable and non-traversable regions quickly and reliably. Integrated in an autonomous robot control architecture, this enables a mobile agent to navigate autonomously in an unstructured outdoor environment.

7. References

[1] J. C. Andersen, M. R. Blas, O. Ravn, N. A. Andersen and M. Blanke: Traversable terrain classification for outdoor autonomous robots using single 2D laser scans. Journal of Integrated Computer-Aided Engineering, vol. 13(3), pp 223-232, 2006

[2] L. Di Stefano, M. Marchionni, S. Mattoccia, “A fast area-based stereo matching algorithm”, Image and Vision Computing 22, pp. 983–1005, 2004.

[3] G. De Cubber, G. Marton. Human Victim Detection. Third International Workshop on Robotics for risky interventions and Environmental Surveillance-Maintenance, RISE’2009- Brussels, Belgium. January 2009.

[4] D. Doroftei, G. De Cubber, E. Colon and Y. Baudoin. Behavior Based Control For An Outdoor Crisis Management Robot. Third International Workshop on Robotics for risky interventions and Environmental Surveillance-Maintenance, RISE’2009- Brussels, Belgium. January 2009

[5] D.B. Gennery. Traversability analysis and path planning for a planetary rover. Autonomous Robots, vol 6, no 2, pp. 131-146, 1999

[6] M. Happold, M. Ollis, N. Johnson, Enhancing supervised terrain classification with predictive unsupervised learning, in: Proc. of Robotics: Science and Systems, 2006

[7] D. Kim, S. Oh, and J. M. Rehg, "Traversability Classification for UGV Navigation: A Comparison of Patch and Superpixel Representations," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 07)*, San Diego, CA, Oct 2007.

[8] D. Kim, J. Sun, S. M. Oh, J. M. Rehg, and A. Bobick, Traversability Classification Using Unsupervised On-Line Visual Learning for Outdoor Robot Navigation, IEEE Intl. Conf. on Robotics and Automation (ICRA 06), Orlando, FL, May 2006.

[9] R. Labayrade, D. Aubert, J. P. Tarel, ”Real Time Obstacle Detection on Non Flat Road Geometry through V-Disparity Representation”, IEEE Intelligent Vehicles Symposium, Versailles, June 2002.

[10] D. Langer. A Behavior-Based system for off-road navigation. IEEE Transactions on Robotics and Automation, vol 10, no 6, pp 776-783, 1994

[11] K. Muhlmann, D. Maier, J. Hesser, R. Manner, “Calculating Dense Disparity Maps from Color Stereo Images, an Efficient Implementation”, IJCV 47, 1/2/3, pp. 79–88, 2002.

[12] L. Nalpantidis, G.C. Sirakoulis, A. Gasteratos, "A Dense Stereo Correspondence Algorithm for Hardware Implementation with Enhanced Disparity Selection", 5th Hellenic Conference on Artificial Intelligence 2008, 2-4 October 2008, Syros, Greece.

[13] L. Nalpantidis, G.C. Sirakoulis, A. Gasteratos, “Review of Stereo Vision Algorithms: from Software to Hardware”, International Journal of Optomechatronics, 2:435-462, 2008.

[14] H. Schäfer, M. Proetzsch, K. Berns, Stereo-Vision-Based Obstacle Avoidance in Rough Outdoor Terrain International Symposium on Motor Control and Robotics – 2005

[15] D. Scharstein, R. Szeliski, “A taxonomy and evaluation of dense two frame stereo correspondence algorithms”, IJCV 7, 1/3, 2002.

[16] H. Seraji. New traversability indices and traversability grid for integrated sensor/map-based navigation. Journal of Robotic Systems, vol 20, no 3, pp. 121-134, 2003

[17] S. Singh et al. Recent Progress in local and global traversability for planetary rovers. Proc. IEEE International Conference on Robotics and Automation, 2000, pp. 1194-1200

[18] M. O. Shneier, W. P. Shackleford, T. H. Hong, T. Y. Chang, "Performance Evaluation of a Terrain Traversability LearningAlgorithm in The DARPA LAGR Program", Proceedings of the Performance Metrics for Intelligent Systems Workshop| 2006 |PerMIS, Performance Metrics for Intelligent Systems (PerMIS) Workshop 2006, Gaithersburg, MD, USA, 08/21/2006 to08/23/2006, (31-Jan-2007)

[19] O. Schreer, "Stereo vision-based navigation in unknown indoor environment", 5th European Conference on Computer Vision, I: 203-217, 1998.

[20] S. Thrun, M. Montemerlo, H. Dahlkamp, D. Stavens, A. Aron, J. Diebel, P. Fong, J. Gale, M. Halpenny, G. Hoffmann, K. Lau, C. Oakley, M. Palatucci, V. Pratt, P. Stang, S. Strohband, C. Dupont, L. E. Jendrossek, C. Koelen, C. Markey, C. Rummel, J. van Niekerk, E. Jensen, P. Alessandrini, G. Bradski, B. Davies, S. Ettinger, A. Kaehler, A. Nefian, and P. Mahoney, "Stanley: The robot that won the darpa grand challenge: Research articles," *J. Robot. Syst.*, vol. 23, no. 9, pp. 661-692, 2006

[21] I. Ulrich and I. Nourbakhsh, "Appearance-Based Obstacle Detection with Monocular Color Vision," Proceedings of AAAI 2000, 2000.

[22] K.J. Yoon, I.S. Kweon, “Adaptive Support-Weight Approach for Correspondence Search”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 4, April 2006.

[23] S. Yoon, S.K. Park, S. Kang, Y.K. Kwak, “Fast correlation-based stereo matching with the reduction of systematic errors”, Pattern Recognition Letters 26, pp. 2221–2231, 2005.

[24] C. Zach, K. Karner, H. Bischof. 2004. “Hierarchical disparity estimation with programmable 3D Hardware”, Proceedings of International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision :275–282.