Obstacle Detection from Uncalibrated Cameras

Nadia BAHA TOUZENE, Slimane LARABI
LRIA, Computer Science Department,
University of Science and Technology – Houari Boumediene
Algiers, Algeria
nbahatouzene@usthb.dz, slarabi@usthb.dz

Abstract

In this paper, we propose a new approach for obstacle detection based on the analysis of images taken by uncalibrated stereo rig. This system can be divided in to two main stages. The first one deals with computing the fundamental matrix from the matching between points of interest in order to compute a dense disparity map. Whereas the second one presents a very simple and faster method for obstacle detection, by using the segmentation image. Indeed, the combination of the segmented image and the disparity map are going to permit us to extract the vertical 3D segments that will indicate us the presence of obstacles in the scene. This approach allows us to detect several numbers of obstacles of varied shapes and sizes. This obstacle detection stage can be viewed as the first stage of a free space estimator which can be implemented in a mobile robot.

1. Introduction

One of the classical research problems in the field of computer vision is that of stereo, i.e., reconstruction of the 3D geometry of a scene from a pair of left and right 2D images. Stereo vision has a wide range of applications, including video coding, view synthesis, object recognition and safe navigation.

Obstacle detection [2-5,11,17,18] is one of the fundamental problems of mobile robotics. In order to navigate in the world, it is necessary to detect those portions of the world that are dangerous or impossible to traverse. Detection of the obstacles implies directly or indirectly, the use of some kind of 3D information, and this is the reason why the active sensors, such as laser or radar, are the prime choice of industry. However, the use of a high resolution, high accuracy stereovision algorithm provides comparable results in 3D estimation, while delivering a larger amount of data, thus making the grouping and tracking task easier, and allowing a detection of free space in the indoor environment.

The stereovision –based approaches [7] have the advantage of directly measuring the 3D coordinates of an image. The main constraints concerning stereovision applications are to minimize the calibration and stereo-matching errors in order to increase the measurement accuracy and to reduce the complexity of stereo correlation process.

Our approach is also based on the construction of the dense disparity map from the pair of rectified images that provides us a good geometric representation of the scene. We also applied a segmentation method to the image by using the Sobel filter.

Our main contribution lies in developing a simple method that combines the result of segmentation step and the disparity map computed for the 3D vertical segment extraction that represent the obstacles of the scene. The advantage of this approach is that it is very fast and permits us to detect a big number of obstacles of varied shapes and sizes.

This paper is organised as follows: section 2 presents the stages of the method proposed. In section 3, experiments on real image and an analysis of the results are presented. Finally, we make concluding remarks.

2. Proposed approach

A In this section, we describe the basic steps of our approach. We show how theses steps can be performed quickly and robustly. The stereo images have been acquired from two CCD cameras mounted on mobile robot. The two cameras are at the same height above the ground.

978-0-7695-3323-0/08 $25.00 © 2008 IEEE
DOI 10.1109/PCI.2008.34

2.1. Detection of the features points and matching

The points to detect in an image must be features points of the image in order to facilitate their matching. For it we used the detector of Harris [8] .Once these points are detected we make the matching. The similarity between two portions of two images of a same scene is computed by using the correlation measure. This measure is calculated between two windows, in general squared and centered on the points to put in correspondence. The principal goal of most correlation based algorithm is to find the corresponding point on which to perform the correlation. The traditional way of obtaining the correlation consists of computing the correlation score for every point in the image by taking a fixed window in the left image and a shifting window in the right image. The second window is moved in the right image along the epipolar line.

For each point, the correlation score is computed. Different correlation scores algorithms are used in stereovision. We have chosen the normalized correlation named ZNCC [9] for its robustness. The value of the correlation score vary between -1 and +1, the value +1 indicates the points are greatly correlate, -1 the opposite. In order to validate the corresponding between points, the matching is done from the left image to the right image and from the right image to the left image and we consider valid match only those which give us the best correlation score. Figure 1 shows the left and right images with features points.

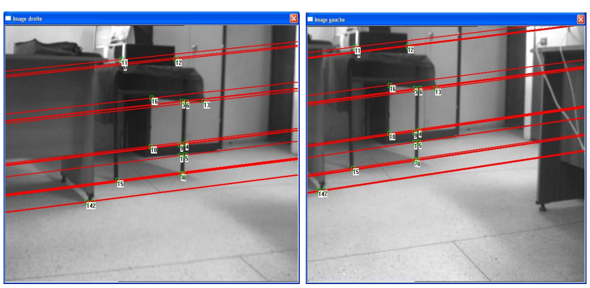


Figure 1 Left and Right images overlaid with features points and the matching.

2.2. Estimation of the fundamental matrix

The fundamental matrix F links the left image to the right image. This means that for each point of the image plane of the left camera, the fundamental matrix 'F' gives us the epipolar line corresponding to it in the image plane of the right camera such as $\mathbf{p^T_r\ F\ p_l = 0}$ where P(x,y,z) is the 3D point and $p_r(x_r,y_r)$ its projection in the right image and $p_l(x_l,y_l)$ its projection in the left image. To estimate the parameters of the fundamental matrix, several methods have been presented in the literature [6,10,13,15]. Most of these methods use a given number of pairs of matched points in the pair images: we find the linear methods and the non linear methods. In our case, we used the non linear methods, although difficult to implement, they offer an enormous advantage because they are robust since they minimize the errors of matching and succeed to a better estimation of the fundamental matrix [15]. Therefore, we used the methods of Gauss - Newton and Levenberg-Marquardts [12].

2.3. Rectification of the stereo images

A constraint that is often considered in stereo vision systems is the epipolar geometry constraint. This constraint involves that any point lying on an epipolar line in one image necessarily corresponds to a point lying on the homologous epipolar line in the other image. Thus, the rectification is an important stage since it allows the use of a simple epipolar geometry in which the epipolar lines are parallel to the lines of images [13,14]. After rectification of the two images, the matched points have necessarily the same ordinate in the two images, and the search of the point of the second image corresponding to a given point of the first image is limited therefore to a one-dimensional search along a horizontal line of the second image situated to the same ordinate, rather than a bi-dimensional search in a region of the second image.

We have chosen to implement the described method in [16], it is based only on the knowledge of the fundamental matrix and the epipolar constraint to rectify a pair of images. It consists in determining two matrixes of rectification one for the left image and the other for the right image. Figure 2 shows an example of the left and right image obtained after rectification step.

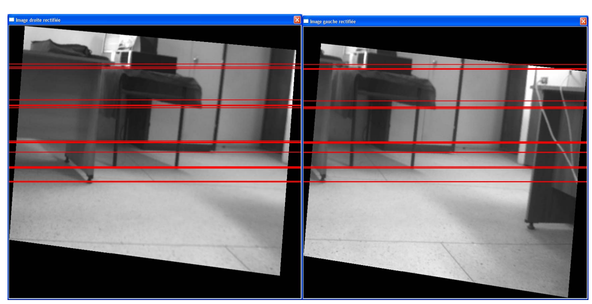


Figure 2 Left and Right images obtained after the rectification step.

2.4. Dense disparity maps estimation from stereo images

The disparity map calculation is expensive both in cost and in time. Our aim is to overcome the limitation by developing a method that with a very low computational cost so that, it can be used for real time applications. The method [1] combines two algorithms: the first one reduces the space of search of the matching points by using one interval of disparity. The second one allows accelerating the process described previously by using a multi-resolution algorithm where the matching is performed at several levels of resolution by using a pyramid structure.

The principle of the computation of the disparity map is for every point of the first image:

1 - Compute the correlation score between windows of the image centered in this point and a window of the same size centered in every point of the other image susceptible to correspond to it, this means that it has the same ordinate since the images are rectified.

2 - Choose as correspondent of this point the one that maximizes the correlation score, and deduct the disparity. It is important to know how to restrict the space of research of the correspondences in the second image for two reasons: Evidently, in the first, the bigger the number of candidates to the correspondence, the longer time of computation. In the second, the bigger the number of candidate, the bigger risks to do false matching. Therefore, we reduce this space of research by the choice of one interval of disparity [Dmin,Dmax], corresponding to a segment of the epipolar line. We determine this interval by measuring the disparity of the features points detected at the first stage of our approach. Dmin corresponds to the smallest disparity and Dmax to the biggest disparity. The method described previously is accelerated by using a pyramidal multi- resolution that has the effect to accelerate the computation and to reduce the interval of disparity [Dmin,Dmax]. A multi-resolution pyramid is a set of images of decreasing resolution (levels), the basis being the origin image (level 1) and the top, the image of lower resolution(level N) .Each level of the multi - resolution pyramid has a resolution of

$A/[n^{(N-1)}]$ by $B/[n^{(N-1)}]$ with:

A: number of columns of the initial image.

B: number of lines of the initial image.

N: number of level (N different from 1).

n: the reduction factor .

The images of reduced resolution are obtained by a reduction by sampling while taking 1 pixel on n (n factor of reduction) in comparison with the previous level (see Figure 3).

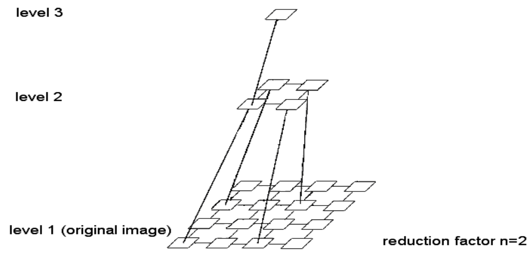
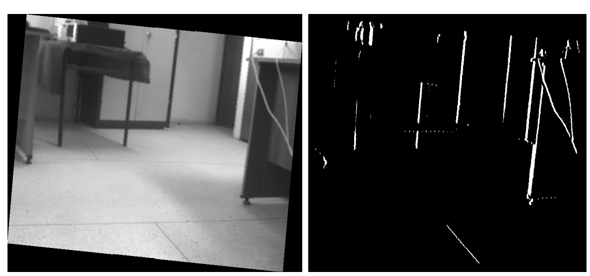


Figure 3 Calculation of the levels of the multi - resolution pyramid.

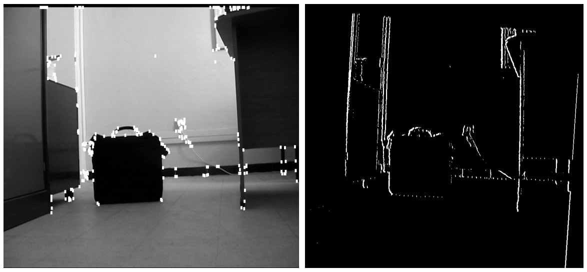
The computation of the disparity map operates, at first, on the superior level which corresponds to the smallest resolution with the ZNCC correlation algorithm. This first step will give us a first disparity map on which we will base to compute the disparity map of the next lower level. For a pixel p(x,y) of level N, we will look for the disparity value "v " that corresponds to it in the reduced image of level N+1, this value "v" allows us to search for the pixel that corresponds to it in a zone centered on the pixel distant of n*v pixels of p at the level N in the other image. As we will go back up progressively in the levels of resolution, we will improve the precision of the results.

2.5. Segmentation

The detection of the contours of an image reduces significantly the quantity of data and eliminates the information that can judge less important, while preserving the structural properties of the image. A several methods of contour detection exist. We have chosen the Sobel filter because it is fast and easy to implement [11] and the results obtained are satisfactory. It detects the horizontal and vertical segment separately on an image in the grey level (See Figure 4(a) and 4(b)**).**



Left image (a) Image segmented



Left image (b) Image segmented

Figure 4 Image obtained after the segmentation stage.

2.6. Obstacles detection

A In the last years, obstacles detection has been a major research topic in computer vision. In fact, in literature, several hypotheses concerning the indoor environments are used in order to facilitate the process [7]. In this paper we assume that most obstacles are of quasi - vertical geometry. From this hypothesis, the presence of a three-dimensional vertical segment in the image can be considered like the presence of an obstacle in the environment of the robot. By using the disparity map previously calculated and the segmentation of the image we can detect the present obstacles in the scene.

2.6.1. Selection of 3D segments. A 3D shape can be approximated by means of one or several 3D straight segments. By using the Sobel Filter, the 3D shapes are decomposed into 3D segments. In this step, in order to select precisely the 3D segments that belong to obstacles, we calculate and threshold the inclination angles of 3D segments with respect to the ground. The inclination angle β of a 3D segment with respect to the ground (See Figure 6) is a pattern that can be used to detect obstacles. Indeed, if the inclination angle of 3D segment tends towards 0, the 3D segment usually belongs to the ground. If this angle tends to 90, then it belongs to obstacles.

In a first step, we compute the vector V of the straight line which contains the 3D segment in the stereoscopic vision system used (Figure 5), and these coordinates constitute the following vector V:

$V = (V_x \quad V_y \quad V_z)$.

Where:

$V_x = p_x * ((m_r * b_l - m_l * b_r) + w/2 * (m_l - m_r))$.

$V_y = p_y * ((b_l - b_r) + h/2 * (m_l - m_r))$,

$V_z = f * (m_l - m_r)$.

With:

m_r, b_l, m_l and b_r are calculated by a least square method.

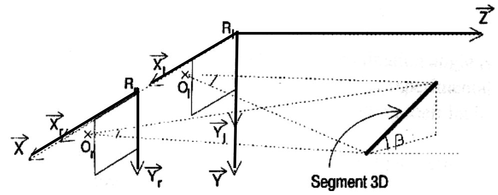


Figure 5 Inclination angle β of 3D segment.

$f_r : x_r \rightarrow m_r * y_r + b_r$, $f_l : x_l \rightarrow m_l * y_l + b_l$

Where:

f_r : the equation of the projection of the 3D segment in the right rectified image.

f_l : the equation of the projection of the 3D segment in the left rectified image.

w: width of the rectified image.

h: height of the rectified image.

f: focal length.

px: width of the pixel of the CCD camera.

py: height of the pixel of the CCD camera.

In a second step, the inclination angle of 3D segment is computed with respect to the ground by using geometry properties. To compute f_r we choose two points that belong to the right projection of the 3D segment. And to compute f_l we look for the correspondents of the two points used in the calculation of f_r by using the disparity map. Indeed, $xl = xr \pm d$, where d is the disparity value computed by our algorithm described in section 2.4.

2.6.2. Algorithm of obstacle detection

1. Computing the V coordinates.
2. If ($V_x = 0$ and $V_z = 0$) then $\beta = \pi/2$

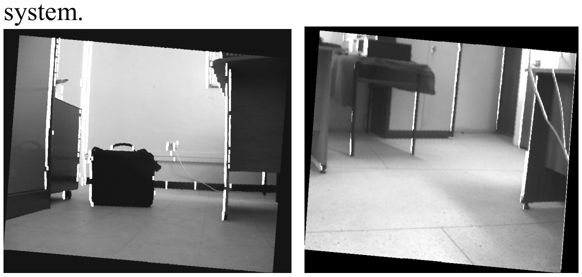
Else $\beta = arctan(-\frac{V_y}{\sqrt{V_x^2+V_z^2}})$

3. If ($\beta \geq$ threshold) then the segment belonging to an obstacle.

In our approach, we are limited to detect the vertical segments. The advantage of this approach is that it is very fast and permits us to detect a big number of obstacles of varied shapes and sizes.

3. Experimental results

To validate the method developed in the previous section, we tested it on two pairs of real images of different scenes. The stereo images have been acquired from two CCD cameras mounted on mobile robot. The two cameras are at the same height above the ground, and the floor can present some motives as shown in Figure 4 (a). The figures 1,2 and 4 illustrate the results obtained at each stage of our system (matching, rectification and segmentation). On the other hand, the computing of the disparity map is made by using the ZNCC correlation with a window of 3x3 sizes. Besides, the disparity map has been calculated with a two levels of pyramidal resolution with factor reduction equal to 2. The Figure 6 shows the experimental results of the extraction of 3D segments of obstacles. The obstacles are represented by the 3D vertical segments. The nearest segment, give us the distance of the obstacle in relation to the stereovision system.



(a) **(b)**

Figure 6: Experimental results of the extraction of 3D segments of obstacles. The obstacles are represented by the 3D vertical segments.

4. Conclusion

A In this paper we have presented an approach for obstacle detection without camera calibration since the approach does not need any information about the configuration of camera. The detection process is based on the rectification of a pair of images and computing a dense disparity map. We also applied a segmentation method to the image by using the filter of Sobel. Our main contribution lies in developing a simple method that combines the result of segmentation step and the disparity map computed for the vertical 3D segment extraction that represents the obstacles of the scene. The advantage of this approach is that it is very fast and permits us to detect a big number of obstacles of varied shapes and sizes. Beside, we will be able to improve this approach by proceeding, in addition to what has been proposed in this paper, to the oblique segment extraction, and make an estimation of the free space in order to construct the free space card which can be implemented in an autonomous mobile robot to allow a safe navigation.

5. References

[1] N.Baha and S.Larabi ,“ Obstacle detection based on dense disparity map”,In 2ème Conference Internationale en Vision Artificielle, CVA2007,Algiers , pp 65-71,2007.

[2] D.Burschka, G.Hager, “Stage classification from dense disparity maps in indoor environments”, In Proc of ICPR Quebec, August 2002.

[3] G.Wang, H.T Tsui, Z. Hu, F. Wu, “Camera calibration and 3D reconstruction from a signal view based on scene constraints”. In Image and Vision Computing pp 311-323, 2005.

[4] Q.T Luong, J.Weber, D.Koller, J.Malik “An Integrated Stereo-Based Approach to Automatic Vehicle Guidance” , in Proc. of the 5th Int Conference on Computer Vision (ICCV-95) pp 52-57, USA 1995.

[5] Z.Zang, R.Weiss, A.R Hansoon, “Obstacle detection based one qualitative and Quantitative 3D reconstruction”, IEEE Tran on pattern analysis and machine Intelligence, pp 15-26, Janv 1997.

[6] A Bartoli, P Sturm,”Non linear estimation of the fundamental matrix with minimal parameters”, In IEEE Trans On Pattern Analysis and Machine Intelligence, Vol 26, no4, April 2004.

[7] N.Simond “ Obstacle Detection from IPM and super-Homography” in IEEE IROS International Conference on Robot and System, Oct 2007.

[8] C.J Harris and M.stephens, “Acombined corner and edge detector”, In Proc 4th Alvey Vision Conf,Manchester, pp 147-151, 1998.

[9] L.Zhong-Dan “Methodes robustes en vision : application aux appariements visuels”. Thèse de Doctorat de L’INP de Grenoble 1997.

[10] B.Boufama, A.Abed, “Three-dimensional structures calculation: achieving accuracy without calibration”. In Image and Vision Computing pp 1039-1049, March 2004.

[11] *H*. Maitre . “Détection des contours dans les images” 2001.

[12] Levenberg *“*A method for solution of certain non-linear problems in least squares”. Quarterly of applied Mathematics, 1944.

[13] Z.Zhang, C.Loop “Computing Rectifying Homograhies for stereo Vision”, IEEE Conf on Computer Vision and Pattern Recognition pp 125-131, Fort Collins,USA June 99.

[14] R.Hartley “Theory and practice of projective rectification”, 1999.

[15] Q.Luong, O.Faugeras, “The Fundamental Martix: Theory, Algorithms, and stability Analysis*”, in* International journal of Computer Vision, 17, pp 43-75, 1996.

[16] F.Devernay. “Vision stéréoscopique et propriétés différentielles des surfaces ”, Thèse présentée pour obtenir le grade de Docteur de l’Ecole Polytechnique Février 1997.

[17] R.Okada, K.Onouchi,”Obstacle detection using projective invariant and vanishing lines”, IEEE Int Conf on Computer Vision France Octobre 2003.

[18] S.Nedevschi,R.schmidt,T.Graf, R.Danescu, “High Accuracy Stereo Vision System for Far Distance obstacle Detection”, IEEE Intelligent Vehicles Symposium, Parma, Italy, June 2004.