Obstacle detection algorithm based on stereoscopic images

A navigation aid system for the visually impaired

ELACHHAB Adil, MIKOU Mohammed

University Hassan I, Faculty of Sciences and Techniques
Laboratory: Analysis of Systems and Information Processing
BP 577, 26000 Settat, MOROCCO

Abstract— This article presents an algorithm for stereoscopic image processing taken by the same camera in a real environment and considering two shifted captures. The aim of this treatment is to extract two important informational data about the environment: the distance separating the obstacles from the camera and their sizes. The depth calculation of an obstacle is performed by exploiting the disparity map between two considered stereoscopic images. This disparity is estimated by using the Block Matching program implemented in Matlab. In order to obtain an accurate and fast result, improvements are added to this estimation by combining "sub-pixel accuracy", "dynamic programming" and "the pyramid image technique". Finally, the determined disparity, associated with the camera calibration data, allowed us to obtain a precise estimation of the depth. The obstacle size is obtained by using the established calibration curve, giving the proportionality between the surface in pixels² of the image of object, and its real surface in m², according to the depth of the object. The tests performed in a real environment based on the elaborated algorithm give the results about the obstacle depth and its size. Tests results are completely similar to the experimental data.

Index Terms— visual impairment, stereoscopic vision, sub-pixel estimation, dynamic programming, pyramidal image.

1 INTRODUCTION

THE vision allows us to perceive and interpret the world around us. For the majority of people with visual impairment, move independently, safely and effectively in unknown places can be a very difficult task to accomplish. The artificial vision is intended to reproduce some functionalities of human's vision through the image analysis [1-4]. Thus, many research groups have proposed devices to facilitate the navigation [5-7]. The most relevant problem for blind people is to evolve in their environment without coming up against unexpected obstacles. The obstacle detection is therefore one of the main problems to resolve and ensure a secured navigation.

The obstacle detection is one of the most active research domains in computer vision, through its wide range of applications, which includes the autonomous navigation, robot and conduit systems [8-10]. The most important objective is to measure accurately the distances relative to obstacles and to provide an estimation of the obstacle size. The existing solutions cannot produce results with sufficient accuracy in real time. In this light, an assistance system can be implemented. This system exploits the stereoscopic images information acquired by two similar cameras, to precisely extract the object volume and the distance separating it from the cameras. Once obtained, this information can be transmitted in real time to the visually impaired or the blind as audible or vibrating signal.

The matching of stereoscopic images, which means to find for each point on the left image, the corresponding point in the right image, constitutes one of the most studied subjects in the field of computer vision. So, to solve this problem, many methods were suggested [11, 12]. To determine the pairs of conjugate points of the two stereoscopic images, it is necessary to evaluate the similarity between the points by performing a comparison of pixels intensity in the two images. This matching allows to establish a disparity map between the two stereoscopic images.

However, the determination of a disparity map often requires important calculation time and precision that are constraining for the entire system. The majority of research efforts have been concentrated on the accurate estimation of the disparity in order to obtain accurate 3D information [13,14]. A disparity describes the difference in the positions of two corresponding pixels. To get the disparity map, we must solve the correspondence problem for each pixel.

In the previous work [15], we proposed an algorithm to calculate the distance separating the obstacles in a scene relative to the camera position, based on the disparities evaluation between the two stereoscopic images, by exploitation of the Block Matching algorithm implemented in software "Matlab", using the sum of absolute differences (SAD). By applying the sub-pixel precision technique through the measurement of similarity functions based on windows, the disparities were determined with a correct accuracy. The exploitation of dynamic programming has considerably increased the accuracy. This algorithm has improved the accuracy of the estimated disparities, and has also accelerated the treatment process through the implementation of the pyramidal image process. These results previously obtained are insufficient for an application in a real environment and that requires real-time processing, therefore we propose in this paper an improved algo-

ISSN 2229-5518

rithm whose aim is to perform a quick and accurate processing to calculate the depth of an object in a real environment. The advantage of this algorithm compared to the previous one, is that it allows separating the obstacles of the background before the image processing. The Canny filter localizes the largest differences of local intensities and finds the object edges in the image. These edges are necessary to generate the binary mask that allows an entirely automatic selection of the obstacle. The most important features are: 1) an acceptable processing speed, 2) a capacity of minimization false matches, and 3) a correct estimation for large distances cameras-object. In our study, we will also establish calibration curve to evaluate the real dimensions of the objects depending on the dimensions of their images taken by the cameras by considering different distances object-cameras.

2 Methods and Materials

2.1 Methods of processing stereoscopic images

The basic structure of a stereoscopic vision system is described in Fig.1.

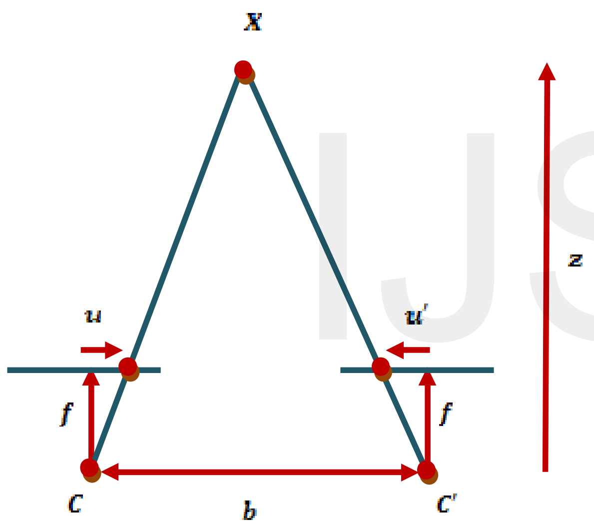

Fig. 1.Geometric model of stereovision system

C (resp. C') : The left camera (resp. right),
f : Focal distance of camera (The same camera is used),
b : Distance between the two positions of the camera,
X : An object point in the scene,
u (resp. u') : The projection of the point X in the left image (resp. right).

Two similar cameras are separated by a distance b, called the baseline. Each camera sees the object but from a slightly different angle. The depth Z, representing the separation distance of the object from the cameras is calculated as shown in the following equation:

$$Z = \frac{b \times f}{d}$$

Where d is the disparity. It measures the distance separating two stereoscopic point images superposed, associated at the same point of an object (d=u-u').

2.2 Presentation of the proposed algorithm

In this section, we explain the global structure of the proposed algorithm. This latest begins by performing a preprocessing of the taken stereoscopic images.

Considering the image in its global shape in the processing provides a better precision, but this increases the calculation complexity, consequently, a two-dimensional binary mask is applied to a block of image data, this block frames the object in the scene that can be seen as an obstacle for the blind or visually impaired.

The obstacle in the foreground of the image is separated by using an edge detection followed by morphological operations. The correspondence research space is therefore reduced by the research around the target obstacle.

Our algorithm has two variables: the size and depth of the obstacle. First of all, the initial disparity map is calculated. Then, the algorithm eliminates noisy disparities by applying the sub-pixel precision and dynamic programming, as well as, the pyramid image to reduce computation complexity. Once the problem of stereoscopic matching is resolved, by exploiting the intrinsic parameters of the camera, the obstacle depth can be estimated.

Finally, the calibration curves provide an estimate of the obstacle size.

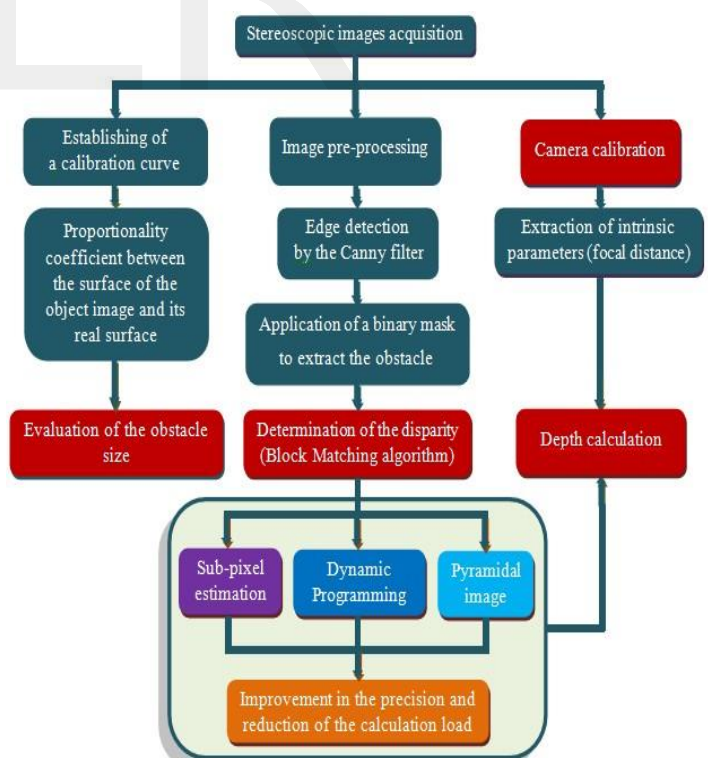

Fig. 2. Schematic diagram of the proposed algorithm

ISSN 2229-5518

The algorithm block diagram of determining the size of an obstacle and its location relative to a blind or a visually impaired is shown in Fig. 2. The proposed algorithm mainly consists of three major steps: 1) The camera calibration, 2) the estimation of the disparity between two stereoscopic images, and 3) the establishment of a calibration curve.

2.2.1 Camera calibration and focal distance determination

The first step is devoted to the determination of the focal distancef. For this, the matrix K of the intrinsic parameters which defines the internal geometry and the camera optical characteristics is obtained by a calibration process using the Camera Calibration Toolbox for Matlab.

$$K = \begin{pmatrix} f_x & s & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

$\{f_x, f_y\}$: The horizontal and vertical focal distances in pixels,

s : The skew that defines the angle between the images rows and columns (often close to $\frac{\pi}{2}$),

$\{u_0, v_0\}$: The image center coordinates.

2.2.2. Accurate disparity estimation for an obstacle

2.2.2.1 Image pre-processing

The stereoscopic images are acquired by a single camera, in order to ensure the invariance of the intrinsic parameters of this latter. The captures of views are done in two times from two points of view slightly apart (65 mm). The sensor data is transmitted to the computer for the analyze.

The proposed algorithm has been coded and tested in a MATLAB environment, on obstacles whose the front face surface varies between 1,1088 m² and 18,36 m². For each obstacle, a series of stereoscopic image pairs is collected, for a camera-object distance ranging from 5 m to 30 m with a step of 5 m.

In the pretreatment, the captured image is resized to 256 x 256 pixels; each pixel has a value of grayscale coded from 0 to 255. The read images are colored, but they have been transformed into monochromatic images to simplify the matching process.

2.2.2.2 Automatic selection algorithm of the obstacle

Faced with the increasing complexity of important calculations and data volumes to be processed, we propose a completely automated algorithm for a fast extraction of an obstacle to its background to replace it on a plain background. The selection procedure retains only the pixels belonging to the obstacle based on a binary mask constructed from the edge detection of an image which accelerates the processing and provides robust and accurate results.

2.2.2.2.1 Edge Detection by the Canny filter

The Canny method implements an estimation of the image gradient using the Sobel filter, followed by a hysteresis thresholding of the gradient modulus [16]. A high threshold and a low threshold are to be defined. All the pixels where the gradient modulus is higher than the first threshold are classified as belonging to the image contours. The pixels having a modulus higher than the low threshold, which are connected to the previous segments, are defined as contour points in the resulting binary image.

2.2.2.2.2 The gradient calculation

The gradient that returns the intensity contours. The used operator allows to calculate the gradient following the X and Y directions, it is composed of two convolution masks whose dimensions are 3×3 :

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

The gradient value at a point is approximated by the formula:

$$|G| = |G_x| + |G_y|$$

The detection results of these contours are represented in Fig. 3.

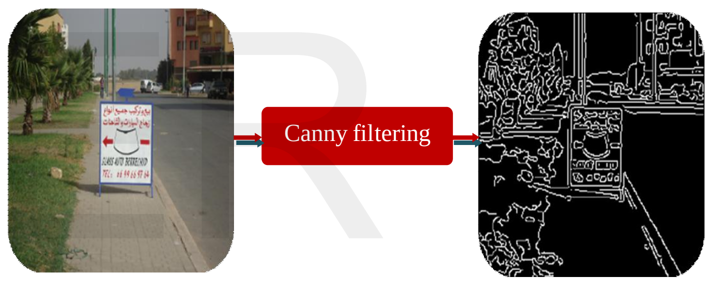

Fig. 3. Edge detection by the Canny filter

Once the edges are detected, the connected components (objects) that have fewer than P pixels, are removed from the map contours.

To extract the object of the image background, first, we compare the grayscale from the center of the image until find a difference higher than three, which means a transition between the object and the background. Afterwards, we register the pixel coordinates of transition for four extremities of the object (rectangular supposed). Then we construct the binary mask from the object coordinates in the image. Finally, we apply the mask on each of the image's component among the three components red green and blue.

The result of this obstacle selection process is shown in Fig. 4.

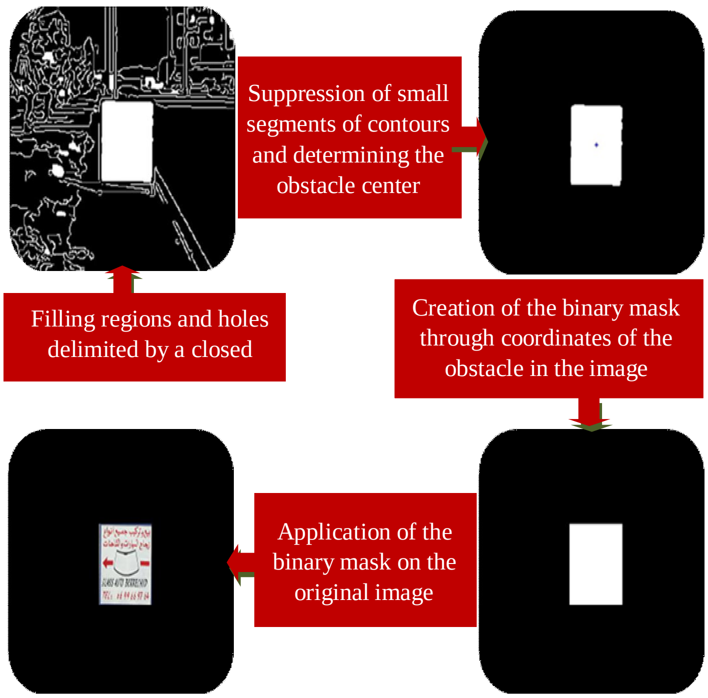

Fig. 4. Execution Steps of the algorithm to extract the object by applying the binary mask

As we can see, the detection is accurate; it allows the complete exclusion of the background.

2.2.2.3 Estimation of the initial disparity map

To determine the initial disparity of the central pixel of size analysis block $15^2 = 15\times15$, the "Block Matching'' algorithm takes this block from the right image, and makes a scan at the same horizontal line in the left image, so as to compare this block with the potential correspondents blocks, found at ± 15 pixels around the pixel location to be matched.
To define the best target block, the algorithm is based on the optimal sum of absolute differences (SAD). The minimal cost of SAD is chosen as the best score and sets the disparity d of the central pixel of the block.

$$SAD(dx,dy)=\sum_{m=x}^{x+N-1}\sum_{n=y}^{y+N-1}\left|I_k(m,n)-I_{k-1}(m+dx,n+dy)\right|$$

2.2.2.4 Improvement of the initial disparity map

2.2.2.4.1 Sub-pixel estimation

The disparity estimations returned by block matching are all integers, so that there are no smooth transitions between regions of different disparities. This causes contouring effects (noisy calculations) that can be eliminated effectively by using the sub-pixel precision [17].
The sub-pixel precision is used to find optimal corresponding points with an accuracy degree less than a pixel, by increasing the harmony of the disparity of a pixel with ±3 disparity values of its neighbors along an image line.

2.2.2.4.2 Dynamic Programming

To have a correct disparity, the dynamic programming algorithm distinguishes the best matching between two sequences among all possible matches, on a range of ± 15 pixels [18]. Whilst the dynamic programming can improve the accuracy of the stereoscopic image, the matching of the basic block is always an expensive operation, and the dynamic programming only adds the load. A solution consists in use the pyramidal image, the speed is five times faster, since the research is reduced to ± 3 pixels.

2.2.2.4.2 Pyramidal image

It was proved that the approach of the pyramidal image in the stereoscopic matching is faster, such as the search range in each level of the pyramidal image is reduced [19]. Effect a reduction in the level of the pyramid in the pair of stereoscopic images, the size of the research space is considerably reduced.

3 Results

In this section, we describe the undertaken experiments to evaluate the performance of the proposed method using pairs of stereoscopic images taken in real conditions and containing an object placed in the scene at different depths. We also project to achieve fast execution of algorithm, what is the condition of any system of obstacle detection for autonomous navigation. Thus, we present the modeling results. These results are compared with experimental measurements to validate the used model.

3.1 Determination of intrinsic parameters of the used camera

The determination of intrinsic parameters of used camera is indispensable for evaluating the depth Z. To determine these parameters, we used a calibration pattern. Ten images taken with different positions of the pattern were performed. The results of processing of these images taken by means of the "Camera Calibration tool" of MATLAB are shown in the table 1.

Focal distance (pixels)	[181,74312 181,61745] ± [1,21845 1,15536]
Point principal (pixels)	[127,50000 127,50000] ± [1,38230 1,29212]
Inclinaison (degrés)	[0,00000] ± [0,00000] => angle of pixel axes = 90°
Distorsion	[-0,22844 0,23824 -0,01315 0,00801 0,00000]± [0,07409 0,08728 0,00946 0,001681 0,00000]

Table 1: Intrinsic parameters of the used camera

The distortion is in particular caused by the optical system (magnifying effect and decentering). The image distortion coefficients (radial and tangential distortions) are stored in a 5×1 vector.

3.2 Determination of the depth "Z"

The algorithm performance for calculating the depth Z, in terms of the accuracy and the calculation speed, was evaluated by using stereoscopic images for real objects provided by the cameras system (binocular distance: 65 mm, focal distance : 181.74312 pixels). To characterize the dispersion of values of the depth Z, we determined the uncertainty from five repeated measures. The obtained results with the proposed approach, compared with experimental values, are reported in the table 2 and illustrated in the figures 5 and 6.

Z experimental (m)	Determined disparity (pixels)	Z calculated (m)
5 ± 0,003	2,350 ± 0,01	5,02 ± 0,048
10 ± 0,003	1,175 ± 0,005	10,05 ± 0,096
15 ± 0,003	0,792 ± 0,001	14,91 ± 0,104
20 ± 0,003	0,588 ± 0,001	20,09 ± 0,124
25 ± 0,003	0,47 ± 0,001	25,13 ± 0,188
30 ± 0,003	0,392 ± 0,001	30,13 ± 0,213

TABLE 2: EXPERIMENTAL AND CALCULATED VALUES OF THE DEPTH Z SEPARATING THE OBJECT TO THE CAMERAS

As shown in the results, the calculated values of the depth Z are quite comparable to the experimental values. The proposed algorithm is able to correctly identify the obstacles position in both for the small and large distances object-camera. It can be also seen that the processing errors obtained are low, which confirms the accuracy of the used algorithm.

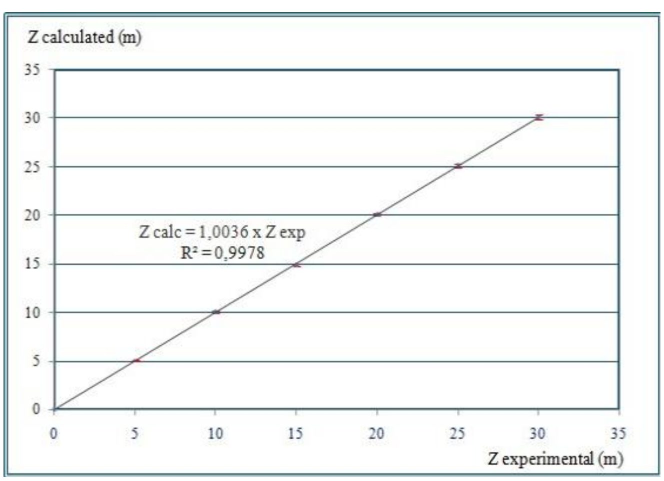

Fig. 5: Comparison of calculated and experimental depths "objet – cameras"

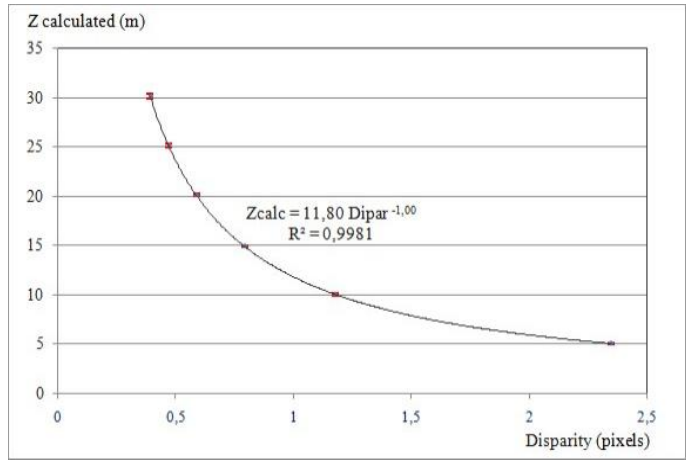

Fig. 6: Variations of calculated depth versus the disparity

Figure 6 shows the variations of calculated depth of objet versus the disparity. The depth values are inversely proportional to the disparities. This result confirms the validity of the proposed method.

3.3 Elaboration of the calibration curve for determining the real size of object

An evaluation of the object size constituting an obstacle for visually impaired people, may constitute useful information about the obstacle importance. For this purpose, we established a calibration procedure to evaluate the real size of the object from the occupied surface by this latter in the image of the scene.

- The first step consists to fix the depth Z at a determined value and calculate the proportionality coefficient α between the surface in pixels² of the object image, and its real surface in m², and this for objects of different sizes. The objective of this step is to verify the independence of the coefficient α relative to the object size. To determine the surface (L*l) in pixels² of the object image, we calculated the difference between its coordinates of the left upper corner and of the right lower corner, which gives the pixels number in the length "L" as well as in the width "l". In our experimental study, we used four objects of different sizes; the obtained results are shown in the table III.
- The second calibration step is to evaluate the proportionality coefficient α, for different values of the depth Z (Experimental or calculated). The obtained calibration curve is shown in figure 7.

The obtained results show firstly that the coefficient α is independent of the object size. The obtained curve of the coefficient α variation as a function of the depth Z is homogeneous, this makes it easy to deduce this coefficient, and consequently, as soon as the determination of the depth Z and the surface s in pixels² of the object image, it easy to calculate the real surface of the object.

Depth Z 5 m	Width (pixels)	Length (pixels)	Surface (s=L*l, pixels²)	Surface (S=L*l, m²)	Proportionality coefficient (α=s/S)	Average $\alpha_{moy} = \frac{1}{4}\sum_{i=1}^{4} \alpha_i$
Object 1	170	178	30 260	18,172	1 665,163	1 666,976± 4,916
Object 2	113	178	20 114	12,048	1 669,489	
Object 3	111	127	14 097	8,448	1 668,679	
Object 4	53	175	9 275	5,572	1 664,573	
10 m						
Object 1	88	90	7 920	18,172	435,826	433,044± 4,563
Object 2	58	90	5 220	12,048	433,267	
Object 3	57	64	3 648	8,448	431,818	
Object 4	27	89	2 403	5,572	431,263	
15 m						
Object 1	58	61	3 538	18,172	194,691	195,975± 3,226
Object 2	39	61	2 379	12,048	197,460	
Object 3	38	44	1 672	8,448	197,917	
Object 4	18	61	1 098	5,572	197,057	
20 m						
Object 1	45	47	2 115	18,172	116,385	116,427± 1,454
Object 2	30	47	1 410	12,048	117,032	
Object 3	29	34	986	8,448	116,714	
Object 4	14	46	644	5,572	115,578	
25 m						
Object 1	36	37	1 332	18,172	73,298	73,389± 0,661
Object 2	24	37	888	12,048	73,705	
Object 3	23	27	621	8,448	73,509	
Object 4	11	37	407	5,572	73,044	
30 m						
Object 1	30	31	930	18,172	51,177	51,513± 0,551
Object 2	20	31	620	12,048	51,461	
Object 3	19	23	437	8,448	51,728	
Object 4	9	32	288	5,572	51,687	

TABLE 3: PROPORTIONALITY COEFFICIENT Α FOR DIFFERENT VALUES OF THE DEPTH Z AND DIFFERENT OBJECT SIZES

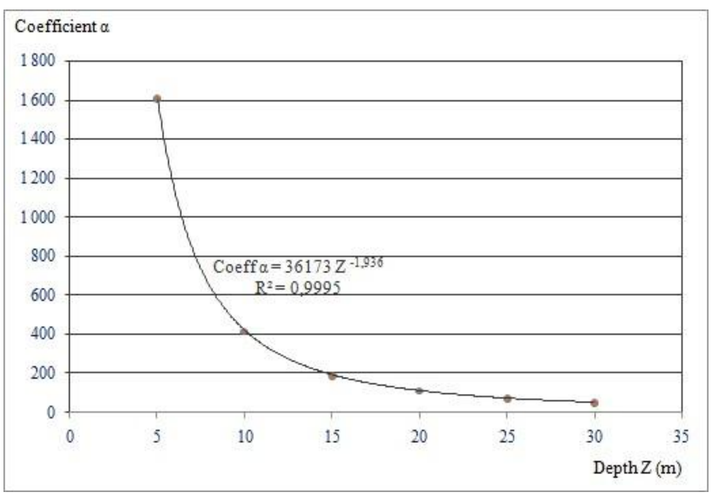

Fig. 7: Calibration curve of the variation of the proportionality coefficient α versus object depth Z

4 CONCLUSION

The main objective of this work is to analyze the technical feasibility of a guidance system based on the stereoscopic vision principle, to help the people with visual impairments to improve their daily autonomy. In this paper, we proposed and evaluated an algorithm of stereoscopic image processing for the evaluation of the size of an obstacle and its location relative to a blind or a visually impaired.

After performing the used camera calibration to define its intrinsic parameters, we estimated and optimized the resultant disparity between taken stereoscopic images of the object by using the dynamic programming algorithm and the pyramidal image, which allowed increasing the processing efficiency in terms of time and precision. The performed processing on the stereoscopic image pairs, associated to the camera intrinsic parameters, made it possible to reach to the values of separation distance camera - object. These values are quite similar to the experimental values

To provide an estimation of the object size constituting an obstacle, we proposed a calibration procedure to determine the variation of the proportionality coefficient α between the surface in pixels² of the object image, and its real surface in m², depending on the depth Z between the camera and the object. The exploitation of this calibration curve allows bringing an appropriate estimation of the object size.

Finally, these results prove that the proposed method is valid in a real environment.

In perspective, it would be interesting to optimize the algorithm to obtain the results of stereoscopic images processing instantly; as well as to develop a translation system of information obtained on the depth and the obstacle size, in the form of a sound or vibrational signal, which will be transmitted to a blind or a visually impaired.

References

[1] R. Gnana Praveen, Roy P. Paily, Blind Navigation Assistance for Visually Impaired based on Local Depth Hypothesis from a Single Image, Procedia Engineering, Volume 64, 2013, Pages 351-360,

[2] Paulo Costa, Hugo Fernandes, Paulo Martins, João Barroso, Leontios J. Hadjileontiadis, Obstacle Detection using Stereo Imaging to Assist the Navigation of Visually Impaired People, Procedia Computer Science, Volume 14, 2012, Pages 83-93.

[3] Yifei Wang, Yuan Gao, Alin Achim, Naim Dahnoun, Robust obstacle detection based on a novel disparity calculation method and G-disparity, Computer Vision and Image Understanding, Volume 123, June 2014, Pages 23-40.

[4] R. Correal, G. Pajares, J.J. Ruz, Automatic expert system for 3D terrain reconstruction based on stereovision and histogram matching, Expert Systems with Applications, Volume 41, Issue 4, partie 2 , Mars 2014, Pages 2043-2051.

[5] Charalampos Tsirmpas, Alexander Rompas, Orsalia Fokou, Dimitris Koutsouris, An indoor navigation system for visually impaired and elderly people based on Radio Frequency Identification (RFID), Information Sciences, Available online 12 August 2014.

[6] H. Fernandes, Vitor Filipe, Paulo Costa, João Barroso, Location based services for the blind supported by RFID technology, Procedia Computer Science, Volume 27, 2014, Pages 2–8.

[7] Johann Weichselbaum, Christian Zinner, Oliver Gebauer, Wolfgang Pree, Accurate 3D-vision-based obstacle detection for an autonomous train, Computers in

ISSN 2229-5518

Industry, Volume 64, Issue 9, December 2013, Pages 1209–1220.

[8] R. Abiyev, D. Ibrahim, B. Erin, Navigation of mobile robots in the presence of obstacles, Advances in Engineering Software, Volume 41, Issues 10–11, October–November 2010, Pages 1179–1186.

[9] Jeongdae Kim, Yongtae Do. Moving Obstacle Avoidance of a Mobile Robot Using a Single Camera, Procedia Engineering, Volume 41, 2012, Pages 911–916.

[10] B. Musleh, A. de la Escalera , J.M. Armingol, Obstacle detection and free spaces in urban environments for advanced driver assistance systems based on stereo vision algorithms implemented in GPU, RevistaIberoamericana de Automática e Informática Industrial RIAI,
Volume 9, Issue 4, October–December 2012, Pages 462–473.

[11] Ju Yong Chang, Kyoung Mu Lee, Sang Uk Lee, Stereo matching using iterative reliable disparity map expansion in the color–spatial–disparity space, Pattern Recognition, Volume 40, Issue 12, December 2007, Pages 3705–3713.

[12] Dongil Han, Byoungmoo Lee, Jae Il Cho, Dae-Hwan Hwang, Real-time object segmentation using disparity map of stereo matching, Applied Mathematics and Computation
Volume 205, Issue 2, 15 November 2008, Pages 770–777.

[13] Sanoj Kumar, Sanjeev Kumar, Nagarajan Sukavanam ,Balasubramanian Raman, Human visual system and segment-based disparity estimation, AEU - International Journal of Electronics and Communications, Volume 67, Issue 5, May 2013, Pages 372–381.

[14] Atsuhiko Banno, Katsushi Ikeuchi, Disparity map refinement and 3D surface smoothing via directed anisotropic diffusion, Computer Vision and Image Understanding, Volume 115, Issue 5, May 2011, Pages 611–619.

[15] Elachhab Adil, Mikou Mohammed, Feasibility Study of a Navigation Aid System for The Visually Impaired, Determination of the depth and size of an obstacle by stereoscopic images processing, International Journal of Engineering Research & Technology (IJERT), Vol. 3 Issue 4, April – 2014.

[16] Canny, J., 1986. A computational approach to edge detection. IEEE Trans. PAMI, 8, 679-698.

[17] Edgar Arce, J.L. Marroquin, High-precision stereo disparity estimation using HMMF models, Image and Vision Computing, Volume 25, Issue 5, 1 May 2007, Pages 623–636.

[18] P.H.S. Torr, A. Criminisi, Dense stereo using pivoted dynamic programming, Image and Vision Computing, Volume 22, Issue 10, 1 September 2004, Pages 795–806.

[19] Sylvie Chambon, Alain Crouzil, Occlusion handling in correlation-based matching, Signal Processing, GRETSI, Volume 24, Issue 6, 2008, Pages 429-446.