

Feasibility Study of a Navigation Aid System for The Visually Impaired

Determination of the depth and size of an obstacle by stereoscopic images processing

Elachhab Adil

University Hassan I, Faculty of Sciences and Techniques
Department of Applied Physics
Laboratory: Analysis of Systems and Information
Processing
BP 577, 26000 Settati, MOROCCO

Mikou Mohammed

University Hassan I, Faculty of Sciences and Techniques
Department of Applied Physics
Laboratory: Analysis of Systems and Information
Processing
BP 577, 26000 Settati, MOROCCO

Abstract—Movement and direction, particularly in unexplored environment remains a major challenge for the visually impaired. For contribute to reducing this limitation, we propose in this work to study and verify the validity of a navigation aid system based on stereoscopic vision. This system consists of two similar cameras allowing taking two stereoscopic images of the same object. The treatment of these two images provides the ability to evaluate the separation distance between the object and camera. For this effect, we propose a processing algorithm of stereoscopic image pair captured by the cameras. In this algorithm, we associate with the block matching, sub-pixel accuracy and dynamic programming to precisely locate the disparity between the two stereoscopic images and optimize the processing time. In fact, the technique of the pyramid image makes a great improvement in processing speed. The data processing stereoscopic images recovered, associated with calibration data of the camera, have yielded an accurate evaluation of the separation distance object - cameras.

In the goal to provide an estimate size of the object being photographed, we performed a calibration curve allowing converting the measured surface in pixels² of image, at the real surface of object in m².

The results obtained from this preliminary study are encouraging and arouse the interest of developing a stereoscopic guide system in a real environment.

Keywords— *visual impairment; stereoscopic vision; sub-pixel estimation; dynamic programming; pyramidal image*

I. INTRODUCTION

The number of people suffering from visual impairment will not stop growing in view of longevity. This handicap threatens the wellbeing of people especially in the present actual world where visual information is becoming increasingly ubiquitous and indispensable.

Thus, to ensure the independence of visually impaired people in the performance of their daily tasks and improve their overall quality of life, different solutions have been implemented or under development. Among these solutions: The bionic eye, it's an implant that is inserted at the retina. It consists of a small microelectronic chip sensitive to light and can send, through several microelectrodes, the signals

corresponding to the image formed at the bottom of the eye directly to the optic nerve, and thence to the sensory areas of the brain, thereby generating a stimulus that simulates the visual information [1,2]. Other method, the ultrasonic guidance based on ultrasonic waves reflection on an obstacle and their return to their starting point by giving real-time information on the direction and the distance of separation from the obstacle [3]. The orientation by satellite is also another method of guiding: A mobile GPS uses a voice synthesizer to convey information to the blind or visually impaired on the environment that surrounds [4,5]

In this context, another computer vision system can be implemented. This system is based on stereoscopic vision. Indeed, the processing of images of an object acquired through two similar cameras, using dedicated mathematical algorithms, allows providing an important information on the object, such as its size 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 under form of sound or vibrational signal.

Generally, the computer vision is a pivot of artificial intelligence that does not seek to replicate the human vision, but to create an algorithmic model using approaches that provide approximate information to those obtained by human vision. The feature extraction from stereoscopic images has been the subject of various researches and has been of great interest in the fields of computer vision and image analysis [6-8]. The Block Matching algorithm remains one of the most appropriate solutions for matching similar blocks between two stereoscopic images [9-11], by means of the sum of absolute differences (SAD) [12]. The concepts of dynamic programming [13,14] and the pyramidal image [15,16] have been widely developed in applications to solve some optimization problems and to accelerate the process of the stereoscopic images.

As part of this work, we propose to develop a method of processing of stereoscopic images taken by the same cameras in order to determine the depth of the separation of object to the cameras. For this purpose, we exploited the Block Matching algorithm implemented in software "Matlab". The method of matching stereoscopic images by "Block Matching" often requires an important computation time and a significant

accuracy. To respond to these imperatives, the dynamic programming has the advantage of considerably increase the accuracy. The method of the pyramidal image, for its part, allows significantly reducing the memory required during processing. It also reduces the number of operations performed and therefore accelerates the processing process. Thus, in our study, we combine these two methods in order to gain in computation time and the accuracy. Finally, the processing result of the stereoscopic images allows determining the disparity between the two images. This disparity is due to the different position of taking pictures. We will also establish calibration curves to evaluate the real dimensions of objects depending on the dimensions of their images taken by the cameras.

II. MATERIALS AND METHODS

A. Technique of taking stereoscopic images

We took the stereoscopic image pairs using a single camera, in order to ensure the invariance of the internal parameters of this latter. The capture of two images of the same scene is taken in two times from two points of view slightly apart. The distance between the two taken views corresponds to average distance between the eyes (65 mm). To prevent any vertical deviation and any rotation during taking pictures, the camera was fixed on a rigid linear support, the two taken of view are parallel and point to the same direction without convergence. The process of taking stereoscopic view is illustrated in the figure below.

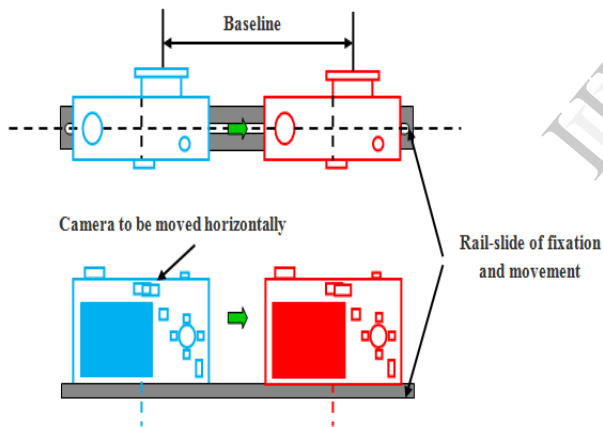


Fig. 1. System of taking stereoscopic view with a single camera

The objects used in this study are parallelepipedic boxes whose the front face surface varies between 4.2 cm² and 414.8 cm². For each object, a series of stereoscopic image pairs is collected for a camera-object distance ranging from 1 m to 4 m with a pitch of 50 cm.

The sensor data is transmitted to the computer for analysis. The images are represented by matrices of pixels. Each pixel has a value of gray scale coded from 0 to 255. The taken images are in color, and converted to gray scale in order to simplify the matching process. The use of full image in the treatment offers better accuracy, but this increases the computation complexity. For this purpose, we have applied two-dimensional binary mask to the image data. This mask

selects a block of image which seems representing an obstacle for the visually impaired.

B. Images processing methods

In this part, we present the implemented processing methods for the extraction of the depth Z , representing the separation distance between the obstacle and the cameras. The geometric model allowing the determination of Z is described in figure 2:

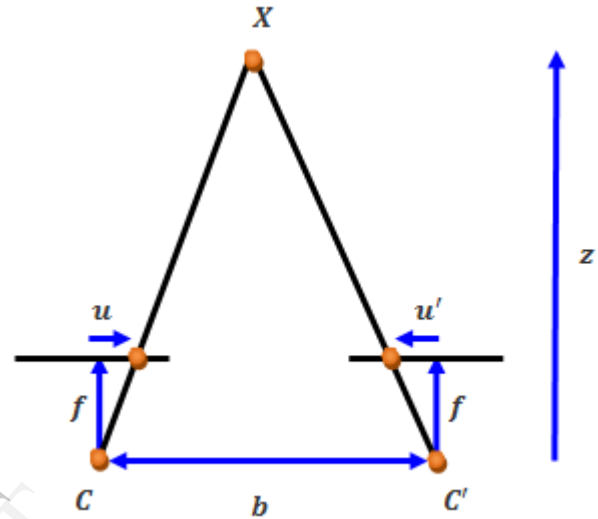


Fig. 2. Geometric model of stereovision system

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).

The depth Z is formulated as follows [17]:

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

Where d is the disparity. It measures the distance separating two stereoscopic point images superposed, associated at the same point of an object ($d = u - u'$). Its value increases as the distance between the object and the stereoscopic camera decreases.

B. 1. Camera calibration and focal distance determination

A way to determine the focal distance f , is to use the matrix K of the intrinsic parameters that provides modeling the internal geometry and the camera optical characteristics. The camera intrinsic model allows the relation between a 3D point and its projected 2D observable in the image plane [18].

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

$\{f_x, f_y\}$: 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.

The matrix K is determined by taking images of a planar checkerboard composed of points P . The images are taken for different positions of the checkerboard relative to the camera view field (translational movements and rotations are performed on the checkerboard for this effect).

The images taken of the checkerboard are processed with the Camera Calibration Toolbox for MATLAB, according to a calibration algorithm that consists of two main phases:

- Initialization parameters, that calculates a closed form solution on the exclusion basis of any lens deformation.
- Nonlinear optimization, which minimizes the reprojection error (within the meaning of least squares) with distortion.

B. 2. Disparity determination

To determine the disparity, the two stereoscopic images are projected onto a same surface area. The mapping of the images onto the area is performed by the Block Matching algorithm which calculates their difference in the position. Figure 3 shows an example of two stereoscopic images superimposed. The discrepancy observed between the two images represents the disparity.

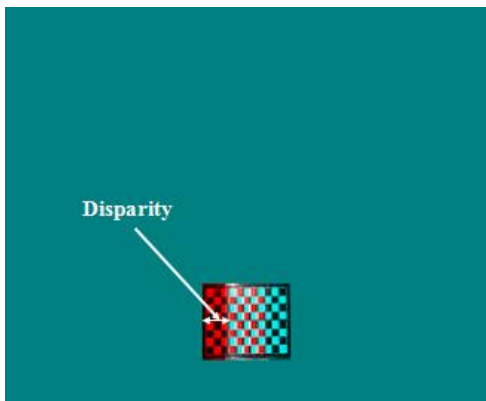


Fig. 3. Superposition of two stereoscopic images

The “Block Matching” algorithm takes a size analysis block $15^2 = 15 \times 15$ (in pixel), centered on the pixel to be matched from the right image, and makes a scan at the same horizontal line in the left image, for compare this block with the potential correspondents blocks, by relying solely on the luminous information (gray scale) (Figure 3).

To define the best target block, the algorithm is based on the optimal sum of absolute differences (SAD) [19], which measures the difference in content between the compared blocks. The best result is obtained for the smallest difference calculated.

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

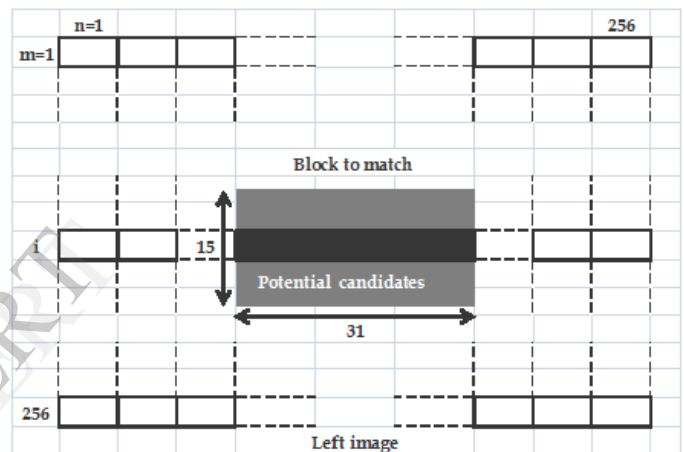
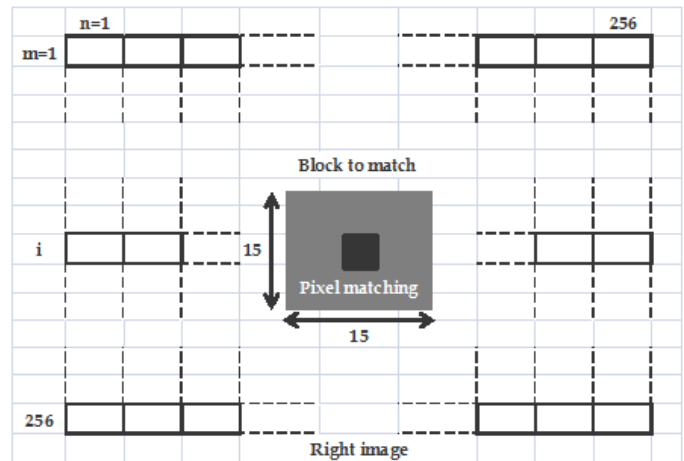


Fig. 4. Matching stereoscopic images by the Block Matching algorithm

A combination between the techniques: Sub-pixel estimation, dynamic programming and Pyramidal image provides improved accuracy and reduces the heaviness of calculation in comparison with the matching by “Block Matching”.

B. 2.1 Sub-pixel estimation

To optimize the disparity map, the sub-pixel estimation technique was introduced. It adds to the location of the minimum cost of the disparity, two neighboring cost values. This will eliminate the contour effects between regions of different disparities.

The “Block Matching” chooses the optimal disparity for each pixel according to its own cost. Still with the aim of improving the quality of the disparity map, a sub-optimal cost is operated and compensated by increasing the harmony of the disparity of a pixel with ± 3 disparity values of its neighbors along an image line.

B. 2.2 Dynamic Programming

The problem of finding the optimal disparity estimates for a pixel row becomes henceforth to find the optimal path of an image side to another. To discern this optimal path, the dynamic programming algorithm distinguishes the best matching between two sequences among all possible matches. The sequences elements to match define the two dimensions of a matrix. The algorithm looks for the optimal path in the matrix. To have a correct disparity, the matching is made on a range of ± 15 pixels. With pyramidal image, the speed is five times faster, since the research is reduced to ± 3 pixels.

B. 2.3 Pyramidal image

The principle consists to create, first of all, a hierarchy of two pyramids of images, each has accuracy four levels, from the finest (256x256) to the grossest (32x32). Then to divide by two the coordinates of primitives in images of lowest resolutions to switch to a grossest resolution. Finally to look for the correspondence along the epipolar line (a set of projection possibilities in the left image as a function of a point of the scene corresponding to a projected in the right image [20]) at the lowest resolution level, while at ascending the left pyramid from 32x32 to 256x256. The point selected by the matching for the image 32x32 corresponds to a region of the image 64x64. The correspondent is the intersection of this region and the epipolar line. And so on until the resolution 256x256 (Figure 4).

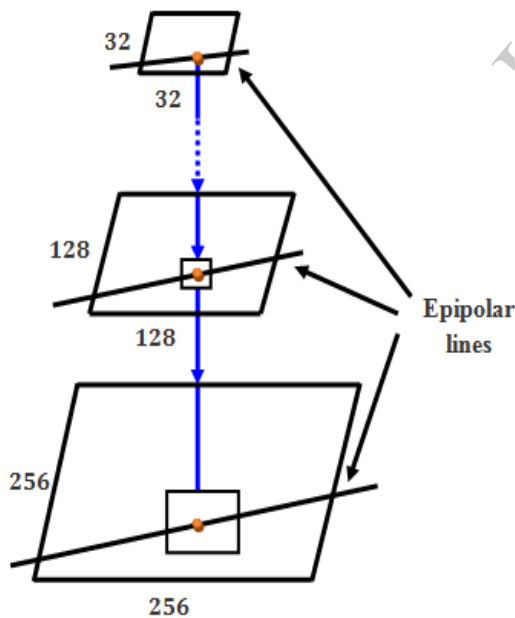


Fig. 5. Searching for the intersection point of the epipolar line and the corresponding region in point of the lowest resolution image

III. RESULTS AND DISCUSSION

In this section, we present the obtained modeling results. These results are compared with experimental measurements in the aim to validate the used model.

A. Calibration results of the used camera

Such as we mentioned in paragraph (II.B. 1), we used a calibration pattern for determining the camera intrinsic parameters. Ten images taken with different positions of the pattern were performed. The results of processing these images taken by means of the “Camera Calibration tool” of MATLAB are shown in Table 1.

Focal distance (pixels)	[186,64555 186,81673] \pm [1,27089 1,30761]
Principal point (pixels)	[127,50000 127,50000] \pm [1,66531 1,82417]
Skew (degrees)	[0,00000] \pm [0,00000] Angle of pixel axes = 90°
Distortion	[-0,26042 0,34733 0,00180 0,00825 0,00000] \pm [0,09299 0,33037 0,00894 0,00792 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 5x1 vector.

B. Depth calculation results

The depth calculation Z of an object on scene relative to the cameras position emanates from the expression of the Z as a function of the cameras separation distance, the focal length f and the disparity. In the following experiments, we test qualitatively the proposed algorithm performance with real stereoscopic pairs taken in normal conditions and containing an object placed in the scene at different locations (paragraph II-A). The obtained results, compared with experimental values, are reported in Table 2 and illustrated in the figures 6 and 7.

Z experimental (m)	0,5 \pm 0,001	Determined disparity (pixels)	25 \pm 0,89	Z calculated (m)	0,48 \pm 0,023
	1 \pm 0,001		12 \pm 0,75		1,02 \pm 0,021
	1,5 \pm 0,001		8 \pm 0,48		1,52 \pm 0,034
	2 \pm 0,001		6 \pm 0,40		2,02 \pm 0,032
	2,5 \pm 0,001		5 \pm 0,32		2,43 \pm 0,046
	3 \pm 0,001		4 \pm 0,28		3,03 \pm 0,035
	3,5 \pm 0,001		3,5 \pm 0,25		3,47 \pm 0,037
	4 \pm 0,001		3 \pm 0,21		4,04 \pm 0,046

Table 2. Experimental and calculated values of the depth Z separating the object to the cameras, obtained for the binocular distance values : 65 mm and the focal distance value 186.55 pixels extracted from the camera calibration.

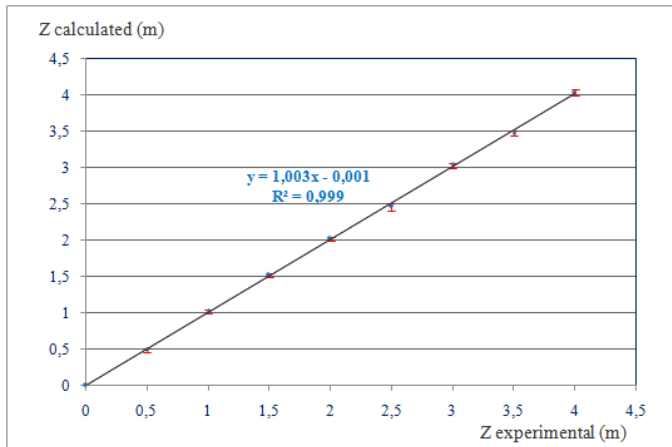


Fig. 6. The calculated depth of the object relative to the camera position depending on the experimental depth

From the obtained result, we find that the depth is correctly calculated for different camera-object distances, in comparison with the experimental depth.

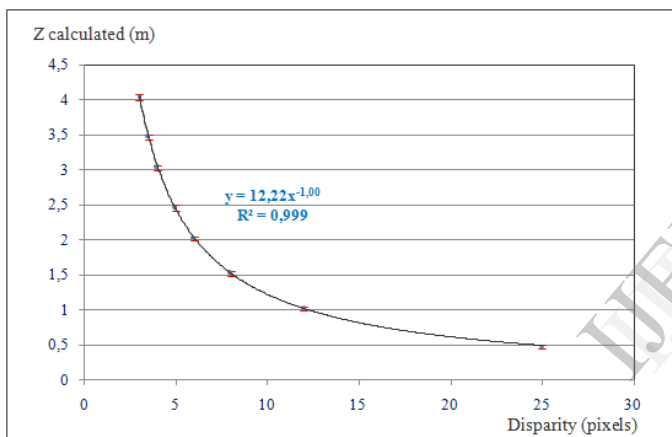


Fig. 7. Calculated depth variation based on the disparity

The figure 7 shows a comparative analysis between the depth estimation and the disparity. The depth values are inversely proportional to disparity, which confirm the validity of the proposed method.

C. Establishing the calibration curves for the object size evaluation in the scene s

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 occupied surface by the object in the scene. Thus, we proposed two calibration curves :

- The first curve is obtained by fixing the depth Z to a determined value and representing the surface variation in pixels² of the object image, depending on the real surface of the object. To determine the surface (L*1) 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 plotted two curves for the values Z = 1 m and Z = 2 m (Fig. 8).

- The second calibration curve is obtained by fixing the object surface, and for different values of the experimental (or calculated) depth Z we evaluated the surface values of the object image in pixels². The calibration curve so obtained (Fig. 9), makes it possible to bring an appropriate evaluation of the object surface at different depths Z.

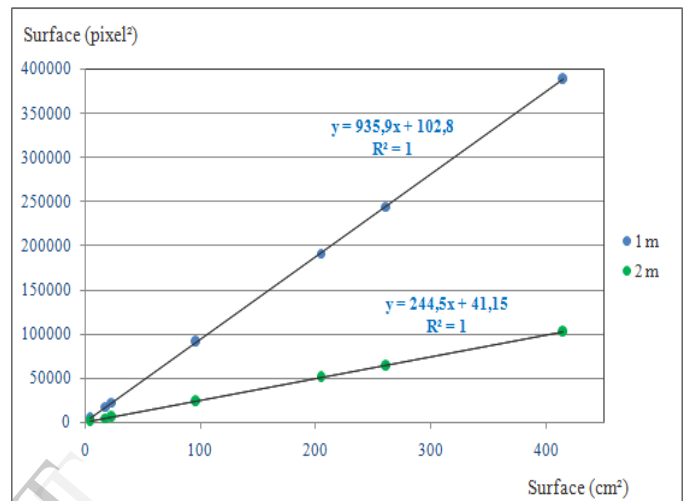


Fig. 8. Calibration curve of the calculated surface variation in pixels² depending on the real surface of the object for the values of Z : 1 m and 2 m

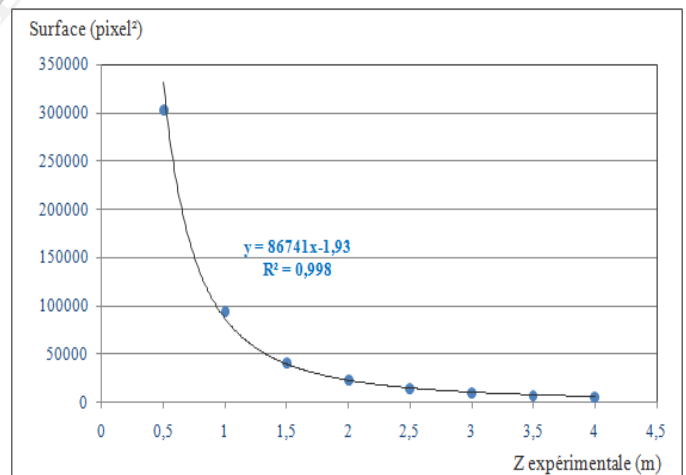


Fig. 9. Calibration curve of the calculated surface variation in pixels² depending on the object depth Z

The calibration result provides an appropriate estimation of the object size in a scene for different depths Z. Thus, this rewarding information allows people with visual impairments to control and avoid the obstacles present in their environment.

IV. CONCLUSIONS AND PERSPECTIVE

The objective of this study is to analyze the technical feasibility of a guidance system based on the stereoscopic

vision principle, to help the people with visual impairments to move freely in their environment. Thus, we proposed and evaluated an algorithm that allows the estimation of the object separation distance constituting an obstacle, relative to similar cameras used for taking the stereoscopic image pairs of the object. After performing the used camera calibration to define its internal parameters, we estimated and optimized the resultant disparity between taken stereoscopic images 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 internal parameters, made it possible to achieve to the values of separation distance camera - object. These values are comparable to the experimental values.

To estimate of the object size constituting an obstacle, we proposed calibration curves of the image surface variation of the determined object in pixels², according to its real surface. Two curves are proposed, the first is obtained by fixing the depth value Z and while varying the object surface, the second is obtained by fixing the object surface and while varying the depth Z. These calibration curves allow bringing an appropriate estimation of the object size.

Finally, these preliminary results seems confirm the validity of the proposed method. In perspective, it would be interesting to check the validity of this model in a real environment.

REFERENCES

- [1] Lotfi B. Merabet, Building the bionic eye: an emerging reality and opportunity, *Progress in Brain Research* 192 (2011) 3–15.
- [2] Katarina Stingl, Karl Ulrich Bartz-Schmidt, Dorothea Besch, Angelika Braun, Anna Bruckmann, Florian Gekeler, Udo Greppmaier, Stephanie Hipp, Gernot Hörtdörfer, Christoph Kernstock, Assen Koitschev, Akos Kusnyerik, Helmut Sachs, Andreas Schatz, Krunoslav T. Stingl, Tobias Peters, Barbara Wilhelm and Eberhart Zrenner Artificial vision with wirelessly powered subretinal electronic implant alpha-IMS, *Proc. R. Soc. B* 2013 280, 20130077, published 20 February 2013.
- [3] Satoshi Hashino, Sho Yamada, An ultrasonic blind guidance system for street crossings, *Computers Helping People with Special Needs* 6180 (2010) 235-238.
- [4] Ponchillia P.E., Rak E.C., Freeland A.L., LaGrow S.J. Accessible GPS: Reorientation and target location among users with visual impairments, *Journal of Visual Impairment and Blindness*, 101 (7) (2007), pp. 389-401.
- [5] Naseer Muhammad, Engr. Qazi Waqar Ali. Design of Intelligent Stick Based on Microcontroller with GPS Using Speech IC, *International Journal of Electrical and Computer Engineering (IJECE)*, Vol.2, No.6, December 2012, pp. 781~784.
- [6] P. Javier Herrera, Gonzalo Pajares, María Guijarro, José J. Ruz, Jesús M. de la Cruz, Combining Support Vector Machines and simulated annealing for stereovision matching with fish eye lenses in forest environments, *Expert Systems with Applications* 38 (2011) 8622–8631.
- [7] Sabri Gurbuz, Erhan Oztop, Naomi Inoue, Model free head pose estimation using stereovision, *Pattern Recognition* 45 (2012) 33–42.
- [8] Hsien-Huang P. Wu, Meng-Tu Lee, Ping-Kuo Weng, Soon-Lin Chen, Epipolar geometry of catadioptric stereo systems with planar mirrors, *Image and Vision Computing* 27 (2009) 1047–1061.
- [9] Fei Yu, Mei Hui, Wei Han, Peng Wang, Li-quan Dong, Yue-jin Zhao, The application of improved block-matching method and block search method for the image motion estimation, *Optics Communications* 283 (2010) 4619–4625.
- [10] Erik Cuevas, Daniel Zaldívar, Marco Pérez-Cisneros, Diego Oliva, Block matching algorithm based on differential evolution for motion estimation, *Engineering Applications of Artificial Intelligence*, 26 (2013) 488–498.
- [11] Changsoo Je, Hyung-Min Park, Optimized hierarchical block matching for fast and accurate image registration, *Signal Processing: Image Communication* 28 (2013) 779–791.
- [12] J. Olivares, J. Hormigo, J. Villalba, I. Benavides, E.L. Zapata, SAD computation based on online arithmetic for motion estimation, *Microprocessors and Microsystems*, 30 (2006) 250–258.
- [13] A. Bensrhair, P. Miché, R. Debrie, Fast and automatic stereo vision matching algorithm based on dynamic programming method, *Pattern Recognition Letters* 17 (1996) 457–466.
- [14] Tingbo Hu, Baojun Qi, Tao Wu, Xin Xu, Hangen He, Stereo matching using weighted dynamic programming on a single-direction four-connected tree, *Computer Vision and Image Understanding* 116 (2012) 908–921.
- [15] Emmanuel TONYE, Alain AKONO, Jean Michel JOLION, Approche Objet Et Pyramidale dans la Classification Non Supervisée des Images de Télédétection, Paris : ORSTOM, 1994, p. 203-217.
- [16] Walid ARIBI, Ali KALFALLAH, Noomène ELKADRI, Leila FARHAT, Wicem SIALA, Jamel DAOUD et Mohamed Salim BOUHLEL, Évaluation de Techniques Pyramidales de Fusion Multimodale (IRM/TEP) d'Images Cérébrales, 5th International Conference: Sciences of Electronic, Technologies of Information and Telecommunications, March 22-26, 2009 – TUNISIA.
- [17] Salvador Gutiérrez, José Luis Marroquín, Robust approach for disparity estimation in stereo vision, *Image and Vision Computing* 22 (2004) 183–195.
- [18] Hui Chen, Hong Yu, Aiqun Long, A New Camera Calibration Algorithm Based on Rotating Object, *RobVis* (2008) 403-411.
- [19] Griselda Saldaña-González and Miguel Arias-Estrada, FPGA Based Acceleration for Image Processing Applications, *Computer and Information Science "Artificial Intelligence Image Processing"*, book edited by Yung-Sheng Chen, ISBN 978-953-307-026-1, Published: December 1, 2009 under CC BY-NC-SA 3.0 license.
- [20] Mingxing Hu, Karen McMenemy, Stuart Ferguson, Gordon Dodds, Baozong Yuan, Epipolar geometry estimation based on evolutionary agents, *Pattern Recognition* 41 (2008) 575–591.