

Fuzzy Free Path Detection from Disparity Maps by Using Least-Squares Fitting to a Plane

Nuria Ortigosa · Samuel Morillas

Received: 24 June 2013 / Accepted: 28 October 2013 / Published online: 24 November 2013
© Springer Science+Business Media Dordrecht 2013

Abstract A method to detect obstacle-free paths in real-time which works as part of a cognitive navigation aid system for visually impaired people is proposed. It is based on the analysis of disparity maps obtained from a stereo vision system which is carried by the blind user. The presented detection method consists of a fuzzy logic system that assigns a certainty to be part of a free path to each group of pixels, depending on the parameters of a planar-model fitting. We also present experimental results on different real outdoor scenarios showing that our method is the most reliable in the sense that it minimizes the false positives rate.

Keywords Assisted navigation · Free path detection · Fuzzy logic

1 Introduction

Autonomous navigation without collisions involves to know the obstacles present in the scene in order to avoid them, or to know the free paths where a vehicle, a robot or a person can walk through.

In order to perform the obstacle detection, there are several references that use different methods, based on laser, ultrasound, infrared or vision systems. For example, in surveillance systems, [1] performs a multi-object detection, and [2] recognizes obstacles by projecting a laser over the ground surface. There are also several contributions based on vision for navigation systems. Some of them are [3], which detects pedestrians by means of stixels computation, [4], which detects and classifies obstacles for automotive applications using stereovision, [5] that detects vehicles and motorcycles using a single-camera, and [6, 7], which perform on-road vehicle detection. Many other contributions also use fuzzy logic in order to provide their results in image analysis, such as [8, 9] and [10], which use fuzzy logic as a motion planner (to decide the next movement of a robot or a vehicle, respectively), whereas fuzzy logic techniques enable to provide uncertainties of objects that have changed their position between two consecutive frames, by means of permanency values in [11]. Particularly, the method proposed in this paper assigns a certainty to be part of a free path, as we will see later on.

In this work we focus on cognitive aid navigation without collisions by means of disparity maps obtained from a stereo vision system. More specifically, we address the obstacle-free path detection problem. Several authors have been working in this area, such as [12] and [13], that use the Sum of Absolute Difference between two stereo images in order to

N. Ortigosa (✉) · S. Morillas
I.U. Matemática Pura y Aplicada,
Universidad Politécnica de Valencia,
Camino de Vera s/n, 46022, Valencia, Spain
e-mail: nuorar@upvnet.upv.es

detect the free space, [14] that analyzes the time-to-impact to objects of the scene based on stereo disparity, [15], that finds the optimum road-obstacle boundary by calculating the disparity space image associated to the stereo pair and finding the maximum edge pixels match under the road environment constraint (turning the stereo constraint into a one-dimensional problem), [16] that differentiates obstacles-free areas by analyzing depths in the scene, or [17], that aids a robot to navigate autonomously and choose the most appropriate path. Another option to detect the obstacle-free space of a scene is to use information from stereo measurements to compute a stochastic occupancy grid (filtered to reduce noise) and transform it from cartesian to polar coordinates, finding the optimal boundary between obstacles and the available free areas in front of a vehicle [18, 19]. More examples are [20], that uses a distance dependent threshold in disparity maps, [21] which is based on the parametric B-splines to estimate the road-surface and [22], that establishes the path to guide a robot by obtaining a digital elevation map, which is used to create a cost map of each possible path. Meanwhile, [23] proposes a direct approach for a three-dimensional road reconstruction using stereo images, [24] avoids obstacles by checking which of the three different parts in which the disparity map is divided presents the lowest disparity value (greater than a determined threshold) to move towards this direction, [25] uses IPM (Inverse Perspective Mapping) to detect free space on highways, and [26] detects obstacles by constructing “v-disparity” images from the associated disparity maps.

Stereo images are often used to estimate depth in scenes by calculating the associated disparity maps. We process a disparity map for each stereo pair of images. So, depending on the possible movement of the cameras, a different number of disparity maps will be calculated for each scene. In case that cameras would be fixed in a determined position, the number of disparity maps associated with that scene would correspond to the frame rate of the cameras per second.

Our algorithm directly uses the disparity pixels to perform a least-squares plane fitting, instead of using them to process occupancy grids, inverse perspective mappings or calculate the time-to-impact to objects of the scene (as references [18, 19, 25] and [14] respectively do). We propose to adjust groups of pixels of disparity maps of the scene to a planar model, since

grey levels of the disparity maps in free paths vary in a linear way along columns and remain nearly constant along rows in free paths. So, depending on the least-squares parameters fitting, we can detect if there are obstacles in an area or not. In particular, we also make use of fuzzy rules in order to obtain a non-crisp detection of free space, since the parameters involved in the algorithm are vague.

This work improves references [27, 28] (which look for linear variations of the associated depth of groups of pixels in obstacle-free areas) by two different ways: 1) it checks whether disparity map pixels fit a plane instead of a linear model, and 2) it uses fuzzy logic to assign certainties to the pixels of being an obstacle-free area in order to deal with partial matching requirements. In addition, it is important to remark that we use real outdoor images which not only present complex lightning conditions and scene ambiguities, but also no constraint on the movement of the cameras (since they are located over the head of the visually impaired user and are constantly in movement, which results on often blurred and deformed objects). Moreover, the whole system must cope with real-time computational requirements over a small and light laptop, without any use of dedicated hardware. Thus, we present our detection method as an alternative of other state-of-the-art references when restrictive conditions in terms of hardware issues, computational cost and constant movement are present when using a prototype.

The rest of the paper is organized as follows. Section 2 presents the contextual framework where the proposed detection algorithm has been developed. The stereovision system and the method used to obtain the disparity maps are described in Section 3. Section 4 details the proposed method to perform obstacle-free paths detection and Section 5 presents the parameters adjustment and experimental results. Finally, Section 6 reports the conclusions.

2 Contextual Framework

This paper presents a real-time free path detection algorithm, using fuzzy logic, in the context of the Cognitive Aid System for Blind and Partially Sighted People (CASBliP) project. The main aim of CASBliP [29] is to provide blind and visually impaired people aid to mobility assistance, interpreting real world

information captured by a pair of stereo cameras and a line CMOS laser sensor, and to sonify this information by means of acoustic maps. 3D-CMOS time-of-flight sensor acquires distance information of the immediate range objects for a line, while stereo vision cameras will determine (apart from the safe pathways where the blind user can walk through) the location, direction and speed per frame of objects relative to the user in the medium to long ranges (of up to 25 meters). Therefore, CMOS sensor alerts will be given priority on obstacles detection for short distances, while stereo cameras will drive the decision making process for objects sensed at depths greater than 5 meters.

Regarding the auditory representation of the scene, CASBlIP uses different sounds for codifying the different objects, in order to improve the quality of the corresponding spatial auditory image. For every detected object, a very short and impulsive spatial sound (a sort of click), which is heard as coming from a certain position in the environment, is used. Thus, free paths are determined by using a specific sound, and they have associated two clicks (which delimitate the obstacle-free area). Depth is indicated by using a higher or lower frequency, depending whether the object is nearer or further. Therefore, the system enables autonomous navigation and identifies potential risks, obstacles and routes.

There exist other references on blind people assistance. An early example is the work of [30], who presented the shape and width of single objects using a matrix of vibrating effectors directly on the skin of a person's back. In [31], scene information was presented to the tongue by electro-vibrating effectors. [32] used both a tonal codification of the grey level and a binaural codification of the azimuthal position for every image pixel, delivering it in a lineal sequential way. Capelle et al. [33] coded every object pixel by means of complex tones which gradually change their tonal, timbric and binaural properties both in height and in azimuth, in order to allow the recognition of simple 2D graphic characters. Recently, [34] introduced a guidance system that detects people, faces and text, and [35] has presented a musical interface to provide obstacle information. In contrast to other references which also relay information to visually impaired people, CASBlIP project is different because of the portability of the system, since all devices work onboard the blind person.

As we have remarked above, the motivation of this work is to detect obstacle-free paths in real-time, in order to assist the blind user in his navigation outdoors. We introduce the innovation of using images of real outdoor scenarios (with complex lighting conditions, scene ambiguities and little constraint on the movement of the user cameras) with real-time computational requirements, too. Thus, we cannot allow to establish any previous requirement, such as having a large planar obstacle-free area in the proximity [36], locating the cameras at a known and fixed height [37], estimate its initial position with respect to the ground [38], or having just one connected obstacle-free area [39], since our cameras will be located over the head of the visually impaired user, and they will be constantly moving around.

3 Stereo Vision Cameras and Disparity Maps

The stereo vision system from the prototype where the detection system is integrated is composed by two aligned Firewire CCD cameras, which provide 240×320 pixel images. This stereo rig will be carried by the end-user. Stereo cameras are located over a helmet or a hat that is worn by the visually impaired user to analyze the scenario that is in front of him. The minimum required distance to the objects to be analyzed is one meter (which corresponds to the necessary distance for an object to be in the field of view of the two cameras). Regarding cameras calibration, it is performed by using a classic chart-based calibration technique [40]. Stereo rectification is also performed in order to set stereo image rows perfectly aligned into a frontal parallel configuration.

Since [41] reviewed many stereo advances in the decade of the 1980s, a substantial progress in stereo research has been made. Last years, many authors have made efforts on developing new matching algorithms and improving the existing ones. [42] presented a taxonomy of dense stereo correspondence algorithms and highlighted the most important features of them. Since then, many new stereo algorithms have been proposed and tested through their website using indoor stereo data sets [43].

Thus, using images acquired with stereo cameras of the prototype, we can obtain and analyze information about depths of objects present in the scene through the disparity map calculation. In addition,

Fig. 1 Examples of stereo images before rectification and disparity maps. *Left and right* images are shown in *first and second columns* respectively and *third column* shows the corresponding disparity maps computed using the reference chosen



according to the purpose of the system where the presented work is integrated, minimizing computational cost without decreasing quality of detection results was a primordial objective. Consequently, we decided to use dense disparity maps as input to our algorithm since other applications of CASBlIP (where our work is integrated) requested them and this did not imply to use any additional hardware. This way, we reuse the disparity maps information, reducing computational cost and hardware requirements, providing results in real-time.

Indeed, as [42] states, dense depth estimation has become an interest research issue in recent years. We have implemented and experimented with different several methods, including dynamic programming [44], sum of squared differences with iterative aggregation [45], and belief propagation [46, 47]. We have found trouble to process many of recent references of stereo matching in the prototype where our work is integrated. Due to the final application (a portable navigation system for blind users) those algorithms which require a special hardware, such as FPGA [48, 49] or a dedicated GPU, are inappropriate for us (our device must be as small and portable as possible). In addition, our system also requires to work at real-time frame-rates. Therefore, references that provide accurate results at expense of raising computational times are not suitable for our work (for example, [50], [47] or [51] result in runtimes of 1-2 seconds for each pair of stereo images). Thus, we have finally chosen the approach described in [44] to estimate disparity maps from each pair of stereo images, since this method is based on dynamic programming and it provides a good trade-off between quality of results and computational efficiency (around 8-10 frames per second on a 1.75 GHz laptop and 1 GB of RAM). In

Fig. 1 we show two examples of stereovision images before rectification and the corresponding disparity maps computed.

In [44], disparity maps are represented as $N \times M$ gray scale images, where lighter areas are associated with nearer regions of the scene. Distance Z and disparity between stereo-images d are related by Eq. 1 [52], where f is the focal length and B is the distance between stereo-cameras (baseline):

$$Z = \frac{f * B}{d} \quad (1)$$

Most of the reported works in obstacle-free areas detection [12]–[21], use stereo images as an input to the detection algorithm, since the use of stereo vision allows to compute disparities between the two images for each frame and perform an accurate detection. The disparity of an image pixel refers the location difference between the pixel in the left image and the corresponding pixel in the right image after both images have been rectified to have a horizontal epipolar geometry. By Eq. 1, the higher the disparity for a pixel is, the lower the distance up to the point represented by this pixel.

4 Proposed Free Path Detection

In this section, the proposed detection method is presented. It is based on the fact that disparity values in obstacle-free areas decrease slightly and linearly from the bottom of the map to the top [28]. In addition, flat zones represent obstacles whose depth is approximately constant. Figures 2 and 3 illustrate this behaviour.

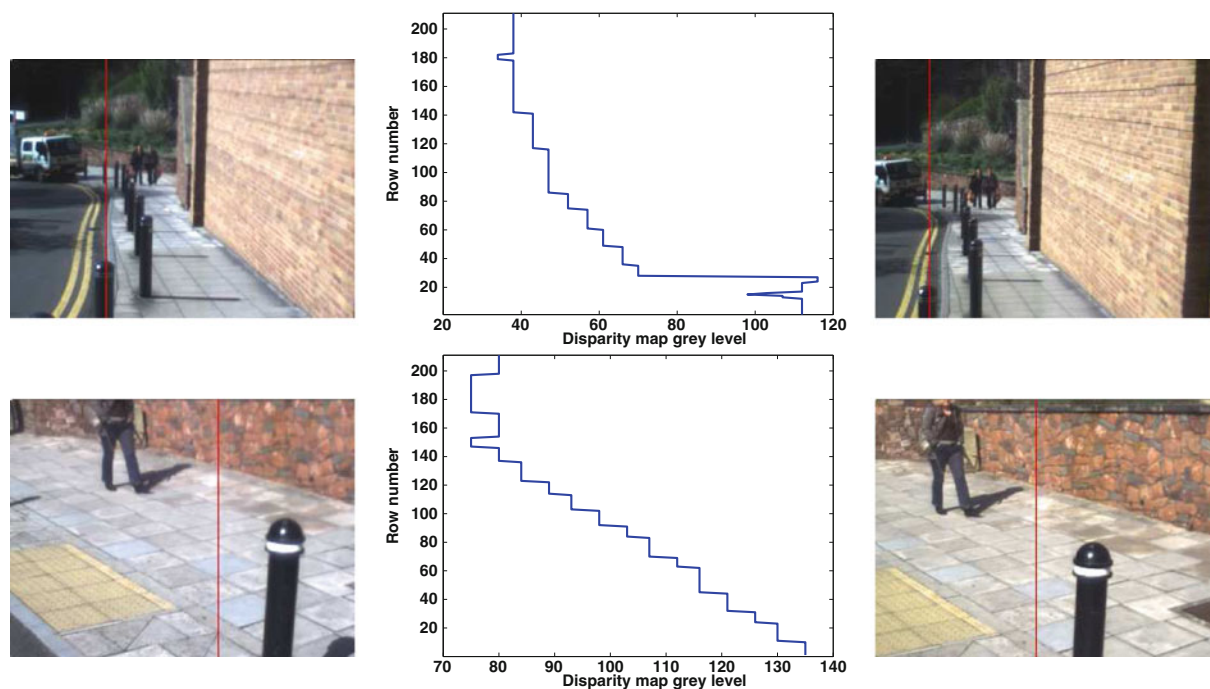


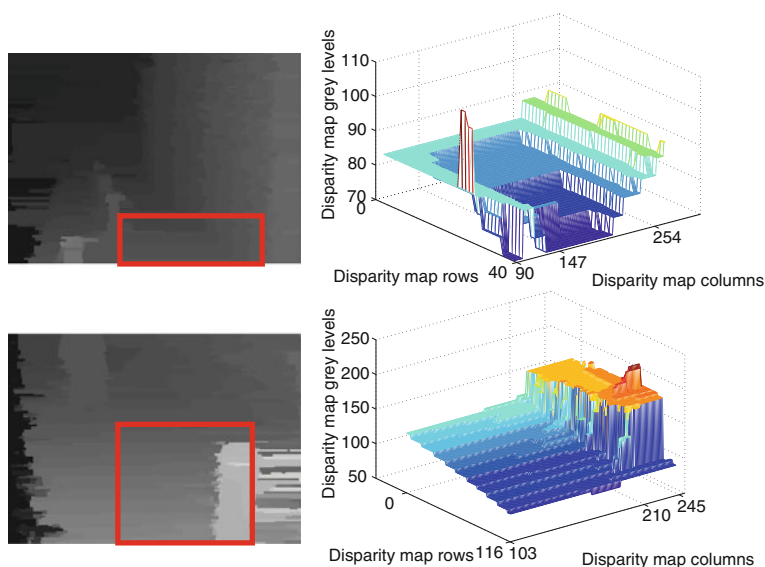
Fig. 2 Two examples to depict the basic principle of the detection method. *First column: Left images* from the stereo-vision system before rectification. *Second column: Disparity map grey levels* for the column images marked in the *left* and *right* stereo images. *Third column: Right stereo images* before being rectified. In obstacle-free areas (rows from 30 to 140 -first example-

and rows from 1 to 150 -second example-) grey levels decrease linearly along the column, from the *bottom* of the image to the *top* of it. Obstacles (the van, the poles and the wall in these examples) are represented with nearly constant grey values, with variations that do not match our linear model, due to the noise of the disparity map

Given a left image I_L and a right image I_R from two stereo cameras, $D(i, j)$ denotes the disparity computed using [44] for pixel in $(row, column)$ location (i, j) .

We are going to adjust to a plane groups of pixels of the disparity map, using a least-squares fitting. So, we divide the disparity map in overlapped windows of size 30×30 pixels. In that window, we first take

Fig. 3 Two examples to show that obstacle-free areas can be adjusted by a plane. *Left: Disparity maps* of the two stereo images shown in Fig. 2. *Right: Disparity map grey levels* shown for the marked areas of the corresponding disparity map. We can observe that disparity values decrease from the *bottom* of the map to the *top* (using a coordinate system positive upwards) whereas obstacles present approximately constant disparity values (columns 201 to 254, and 210 to 245 for the first and the second examples, respectively)



a small central window of 10×10 pixels and then, we take another pixel window of 30×30 around it. This window of disparity values has been weighted by using a gaussian mask so as to the weight for central pixels of the 10×10 window is 1, and the weights for the rest of the 30×30 window decrease as long as they are further away from the central ones. Then, we estimate if disparity values contained in the 30×30 window can be adjusted by using a planar model or not. Next, we move the central 10×10 window to analyze the subsequent group of pixels. Consequently, we implement smooth transitions between the consecutive windows of disparity values (but also maintain higher weights for disparity values of the nearest pixels to the window under analysis). Figure 4 shows the gaussian weights assigned to the 30×30 window.

Moreover, in order to reduce computational cost and to help the whole system to work in real-time, we propose to determine obstacle-free areas by processing only the 25 % last rows of the disparity maps (from row 1 to row $N/4$, using an image coordinate system

whose origin is at the bottom of the image -positive upwards-), since these rows represent the scenario region that the user will first walk through. In fact, we also chose to adjust to a plane each window instead of performing a fitting in a global way since it reduces computational cost, although suitable parameters must be carefully chosen to have a robust method.

The model that provides a least-squares fitting to a plane, over the pixels of the disparity map window is given by:

$$\hat{D}(i, j) = a_{0k} + a_{1k}i + a_{2k}j \quad (2)$$

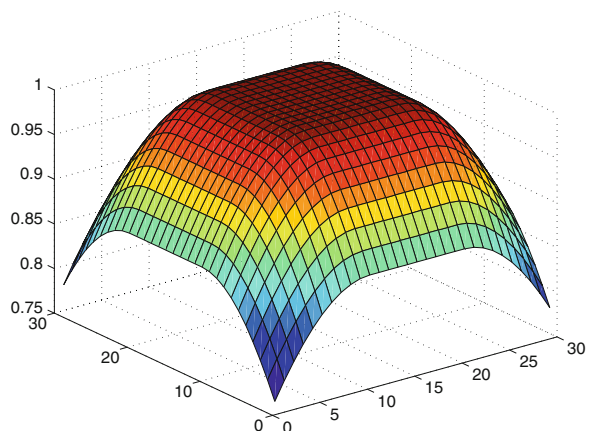
where i and j are the row number and the column number (using an image coordinate system positive upwards) in the processed group of pixels of the disparity map, respectively, $\hat{D}(i, j)$ is the estimated disparity, a_{1k} is the row gradient, a_{2k} is the column gradient of the planar model, a_{0k} is the z-intercept (the point at which the plane intercepts the z-axis), and k is the index of the processed window. Regarding the obtained plane-fitting, its coefficient of determination [53] is defined as

$$I(k) = \frac{\sum_{i=(k-1)*T+1}^{i=(k-1)*T+30} \sum_{j=T*(k-1)+1}^{j=T*(k-1)+30} (\hat{D}(i, j) - \bar{D})^2}{\sum_{i=(k-1)*T+1}^{i=(k-1)*T+30} \sum_{j=T*(k-1)+1}^{j=T*(k-1)+30} (D(i, j) * w(i - (k-1)*T, j - (k-1)*T) - \bar{D})^2}, \quad (3)$$

where $D(i, j) * w(i - (k-1)*T, j - (k-1)*T)$ corresponds to the disparity computed using [44] for pixel located at (i, j) (weighted by the gaussian mask), \bar{D} denotes the mean of the values $D(i, j)$ in the processed window, N is the number of rows of the

disparity map, T is the size of the window (in our case $T = 30$), i can vary between 1 and $N/4$ (since we process only the 25 % last rows of the disparity maps) and j can vary between 1 and M , where M is the number of columns of the disparity map. $I(k)$ measures

Fig. 4 Gaussian mask assigned to the disparity map pixels from the sliding window of the algorithm



the goodness of the fit for the group of weighted pixels from the 30×30 window, and it may take values in $[0, 1]$, where 0 means no correlation and 1 indicates a perfect correlation.

In the following, we will consider the parameters $I(k)$, a_{1k} and a_{2k} and not a_{0k} in the fuzzy rules of the proposed method, since our experimental results showed us that this value is not relevant to assign the certainty of a group of pixels to be a free path. We determine as free paths only those groups of pixels for which we obtain a good adjustment to the planar model. When processing a weighted window, the proposed detection method checks three different relations to perform the detection: it is checked if the value of the coefficient of determination $I(k)$ is *large*, if the gradient a_{1k} indicates a *smooth* variation and if a_{2k} is *small*. a_{1k} and a_{2k} appear performing a dual role in the algorithm since, apart from characterize the fitting of the adjusted plane by means of its coefficient of determination, they behave as relevant features themselves, providing additional information about the horizontal and vertical gradients.

We have considered different possible values of $I(k)$, a_{1k} and a_{2k} in order to make our system more robust in the presence of noisy disparity maps (due to occlusions and wrong matchings) and to provide an additional degree of freedom to our method. Thus, a group of pixels is considered as an obstacle-free area if they fulfill one of these criteria:

1. They adjust well to the plane obtained by least-squares fitting (which implies that $I(k)$ is *large*), the disparities decrease along the columns with a *smooth* negative vertical gradient and they have *small* variations in the horizontal gradient.
2. The coefficient of determination of the obtained fitting is *medium*, their vertical gradient indicates smooth and negative variations, and their horizontal gradient shows that the values hardly vary -it is *very small* - (which means that although the quality of the planar adjustment is poorer, the pixels of the analyzed window correspond to the same obstacle-free area).
3. The coefficient of determination is *medium* (not a high goodness of fit of the planar model), the vertical gradient is *very smooth* but negative (which means that it follows the typical variations of disparity in an obstacle-free area), and the horizontal gradient is *small*.

This reasoning is expressed in vague terms. Since large, smooth and small are linguistic terms, they can be represented by using fuzzy logic [54]. Thus, the membership degree in the fuzzy set *free path* is computed using the following fuzzy rule:

Fuzzy Rule 1: Determining the certainty of $\hat{D}(i, j)$ to be free path

IF “ $I(k)$ is large” AND “ a_{1k} is smooth” AND “ a_{2k} is small”

OR

IF “ $I(k)$ is medium” AND “ a_{1k} is smooth” AND “ a_{2k} is very small”

OR

IF “ $I(k)$ is medium” AND “ a_{1k} is very smooth” AND “ a_{2k} is small”

THEN “the certainty of $\hat{D}(i, j)$ to be free path is high”.

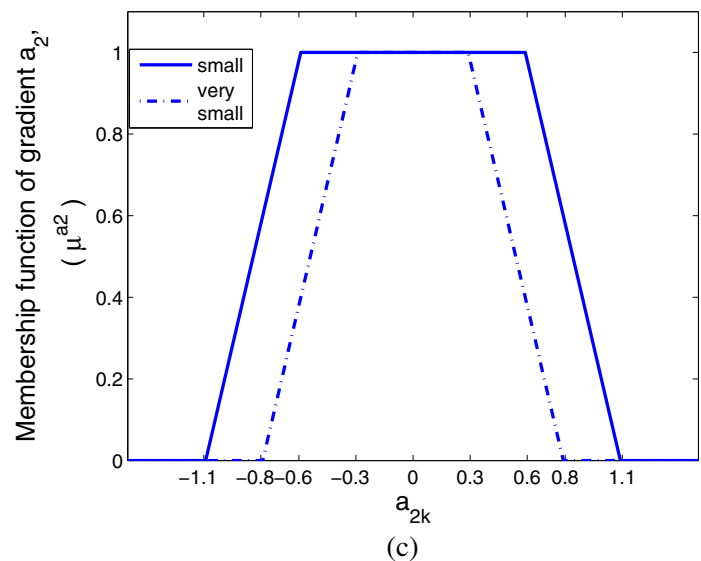
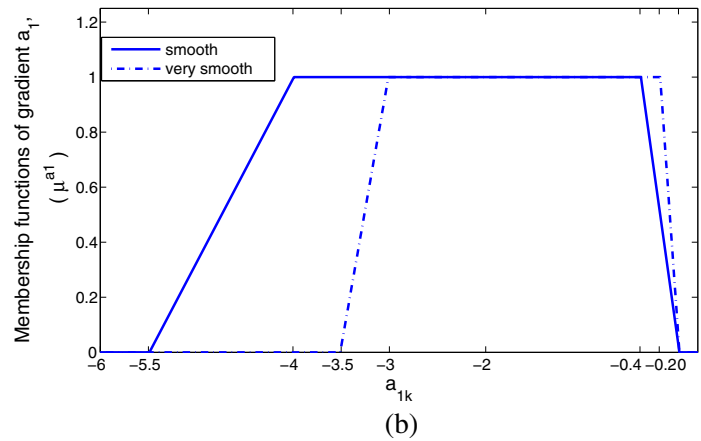
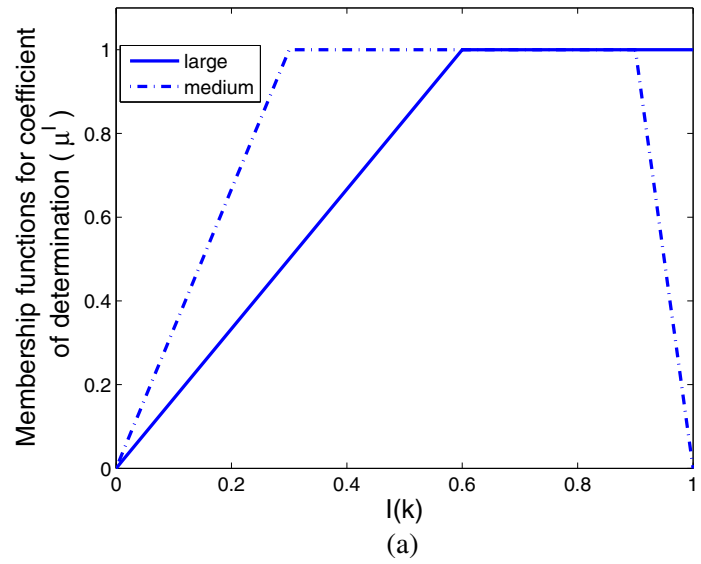
The fuzzy sets in the fuzzy rule can be found as follows. In the fuzzy set *large*, the membership degree should increase as long as coefficient of determination is higher. Furthermore, we want to have large membership degree for a range of negative values of the vertical gradient of the disparity maps (i.e. on the fuzzy set *smooth*) and, in case of being out of that range, the membership degree should decrease slowly. In addition, we want the membership of fuzzy set *small* to be large when the horizontal gradient is near to 0, and then it should decrease slowly.

Thus, whereas there exist numerous types of membership functions, we have chosen a trapezoidal membership function, due to its simplicity [55]. This function is defined as follows:

$$\mu(x) = \begin{cases} 0 & \text{if } x \leq \gamma_1 \\ \frac{1}{\gamma_2 - \gamma_1}x - \frac{\gamma_1}{\gamma_2 - \gamma_1} & \text{if } \gamma_1 < x \leq \gamma_2 \\ 1 & \text{if } \gamma_2 < x \leq \gamma_3 \\ \frac{1}{\gamma_3 - \gamma_4}x - \frac{\gamma_4}{\gamma_3 - \gamma_4} & \text{if } \gamma_3 < x \leq \gamma_4 \\ 0 & \text{if } x > \gamma_4 \end{cases} \quad (4)$$

where the values of the parameters γ_1 , γ_2 , γ_3 and γ_4 vary in the different membership functions. The fuzzy sets *medium*, *very smooth* and *very small* can also be modelled by using the function defined above, just varying the values of the parameters from the ones used for *large*, *smooth* and *small* membership functions, respectively. Figures 5a, b and c show the membership functions for the coefficient of determination and the vertical and horizontal gradients, respectively. All these parameters will be discussed in Section 5.1.

Fig. 5 Membership functions: **(a)** Membership functions *large* and *medium* for the coefficient of determination $I(k)$, **(b)** Membership functions *smooth* and *very smooth* for the vertical gradient a_{1k} , **(c)** Membership functions *small* and *very small* for the horizontal gradient a_{2k}



The fuzzy rule 1 contains six conjunctions and two disjunctions. In fuzzy logic, t-norms are used to represent conjunctions whereas t-conorms are used to represent disjunctions [55–57]. The product triangular norm is used to represent the fuzzy conjunction and the probabilistic sum co-norm to represent the fuzzy disjunction operator. In addition, no *defuzzification* is needed since we aim to obtain a fuzzy obstacle-free degree. Algorithm 1 details the proposed fuzzy detection algorithm.

Algorithm 1: Proposed fuzzy obstacle-free path detection algorithm

1. The disparity map is partitioned into disjoint windows of 10×10 pixels;
 2. **foreach** *disjoint window of 10×10 pixels in the disparity map* **do**
 3. Create a window of 30×30 pixels by taking those ones located around the disjoint window of 10×10 pixels;
 4. Apply the gaussian weights for the current window;
 5. Adjust the parameters of the planar model $\hat{D}(i, j)$ by LMS (2);
 6. Compute the value of $I(k)$ (3);
 7. Determine the certainty for the block of pixels of the processed window 10×10 to be an obstacle-free area using Fuzzy Rule 1;
 8. **end**
-

5 Experimental Results

A set of groundtruth images, in which each pixel was manually marked as part of an obstacle-free area or part of an object were prepared in order to objectively measure the presented detection method performance. This way, we use the mean square error (MSE) of the comparison, pixel by pixel, of those groundtruth images and the obtained detection results to measure the goodness of the fuzzy-detection method proposed.

5.1 Parameters Adjustment

The performance of our method highly depends on the accuracy of the disparity maps. Indeed, complex situations where disparity maps cannot be accurately computed (particularly non-ideally illuminated scenarios) could lead to low performance. That is, bad illumination conditions lead the stereo correspondence algorithms to many false matchings [58], getting worse disparity maps quality. So, in order to make our method as robust as possible and to determine an appropriate parameters adjustment, we have taken a training image set of 35 real outdoor images and we have manually prepared their corresponding

groundtruths. This training image set includes different representative and common complex situations (with different illumination conditions) that a person can find outdoors: vehicles, pedestrians, walls, free paths, poles... Successful navigation in validation trials, carried out using the software integration of the presented method in the prototype of the system, validate the training image set used. Figure 6 shows some of the training images used in the parameters adjustment task.

The parameters involved in the algorithm have been optimized to perform well for a variety of real outdoor images. For this reason, instead of obtaining the optimal parameters for each groundtruth image in the training set, we have found the suboptimal values that maximize the correct detected areas for the whole training image set. The parameters adjustment has been done in an iterative way: first, each parameter is fixed (independently by hand) with an initial value. Second, for each parameter independently we carry out an extensive experimentation about each previous value and we select the best result found (the one which minimizes the mean square error). Third, we evaluate the global performance of the new parameter setting and, if convergence is not reached, we repeat the second step. Figure 7 shows the mean square error increment (in percentage) when the chosen parameters are varied for each membership function.

Thus, we can observe that the parameters adjustment has been found to minimize the mean square error for the whole training set. Finally, we have experimentally found that the parameters should be fixed as follows:

- *Large* membership function for $I(k)$: $\gamma_{1L} = 0$, $\gamma_{2L} = 0.6$, $\gamma_{3L} = 1$, $\gamma_{4L} = 1$
- *Medium* membership function for $I(k)$: $\gamma_{1M} = 0$, $\gamma_{2M} = 0.3$, $\gamma_{3M} = 0.9$, $\gamma_{4M} = 1$
- *Smooth* membership function for a_{1k} : $\gamma_{1S} = -5.5$, $\gamma_{2S} = -4$, $\gamma_{3S} = -0.4$, $\gamma_{4S} = 0$
- *Very smooth* membership function for a_{1k} : $\gamma_{1VS} = -3.5$, $\gamma_{2VS} = -3$, $\gamma_{3VS} = -0.2$, $\gamma_{4VS} = 0$
- *Small* membership function for a_{2k} : $\gamma_{1SM} = -1.1$, $\gamma_{2SM} = -0.6$, $\gamma_{3SM} = 0.6$, $\gamma_{4SM} = 1.1$
- *Very small* membership function for a_{2k} : $\gamma_{1VSM} = -0.8$, $\gamma_{2VSM} = -0.3$, $\gamma_{3VSM} = 0.3$, $\gamma_{4VSM} = 0.8$

Fig. 6 Subset of some of the training images (35 in total) used in the parameters adjustment of the detection method. We can observe that these images include different illumination situations and different kind of obstacles located at different outdoor scenarios

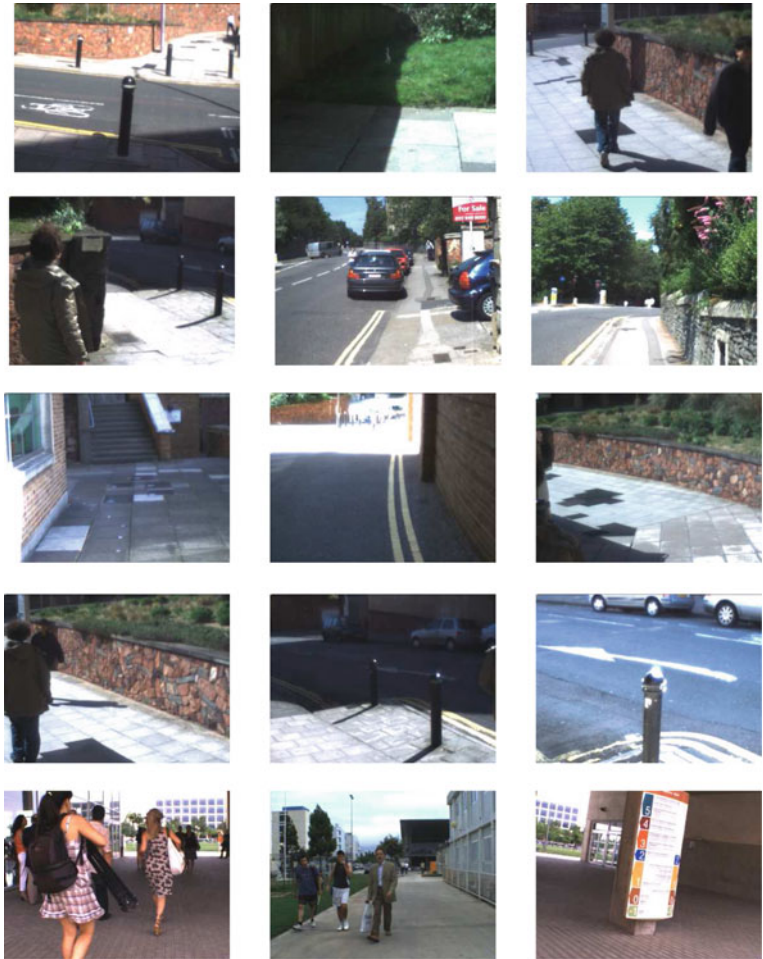


Fig. 7 Mean square error for parameters increment or decrement (in percentage)

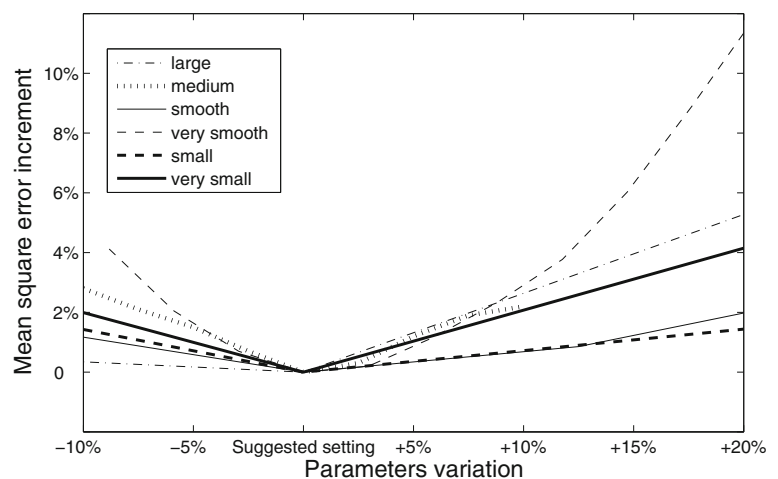


Fig. 8 Processed testing set, where brighter areas show detected free paths. Darker areas indicate a smaller certainty to be part of an obstacle-free area

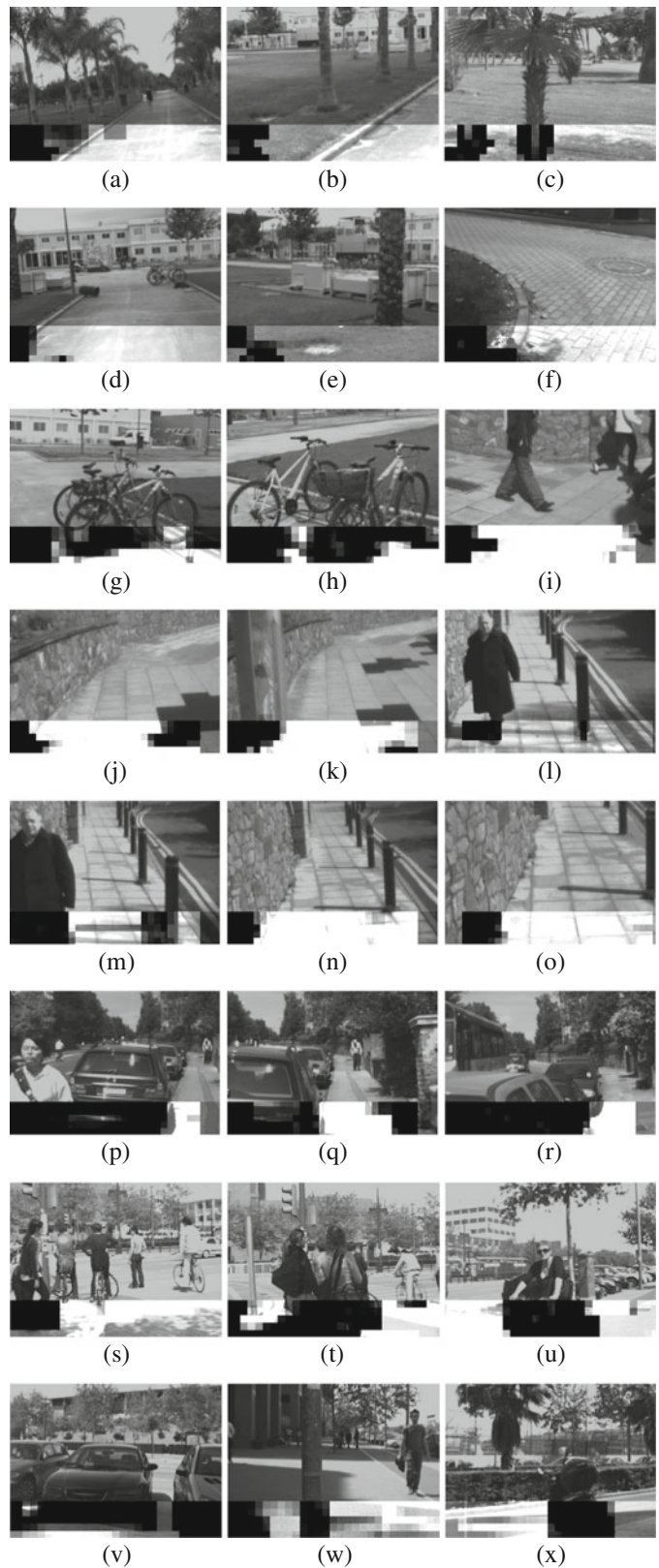
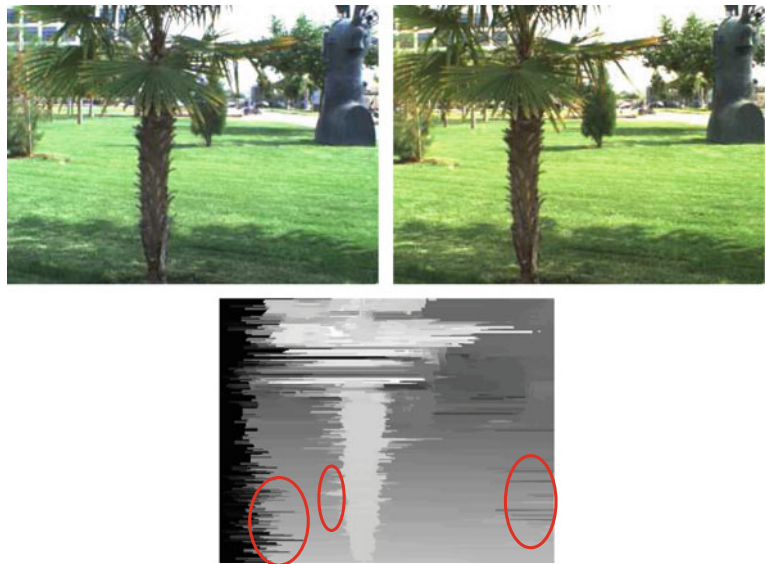


Fig. 9 Original stereo images (*left* and *right*) and their corresponding disparity map with noise. In *red* are marked the zones where we propose to reduce noise



5.2 Algorithm Performance

In this section we assess the performance of the obstacle-free area detection method, by using the previous settings and a set of test images, different from the training image set used for the parameters adjustment (Fig. 6). Figure 8 are examples of the algorithm's output (on original images, for the sake of ease of visualization). Pictures have been created by superimposing the certainties to be part of an obstacle-free area (obtained after processing the disparity maps with the proposed method), depicted them using brightness variations. Brighter areas indicate that their certainty to be a free path is higher than the darker ones. Detection results and figures from Table 1 confirm that the algorithm works well and detects free paths properly.

However, we can observe that there are images where results are not completely accurate. This occurs due to errors in the matching process when obtaining the dense disparity map. Usually, they correspond to

corners of the image (Fig. 8a-b), missing of texture for ground or objects (Fig. 8d, j) or pixels whose corresponding stereo match is difficult to find (Fig. 8e-f in the grass). In addition, this problem significantly increases when, as in our case, the stereo cameras pair is “worn” by a person, whereas most vision systems are fixed or work under controlled movement. Despite this, the presented algorithm always provides the wider obstacle-free area to walk through, so the method performs appropriately for the purpose of the project in which it has to be integrated, in order to avoid collisions of the end user.

Furthermore, matching errors are present in disparity maps, which can be considered as noise (Fig. 9). In order to increase robustness of the detection method against noise, we propose to remove flat black zones (which usually correspond to areas with lack of texture or occlusions and associated disparity is 0) and also remove pixels whose associated disparity varies in an abrupt way repeatedly along the column. We consider

Table 1 Mean square error of the fuzzy detection method for the testing image set, compared with the crisp-linear method presented in [28], using disparity maps with noise and denoised. Columns fifth and sixth show performances for two recent references

	Fuzzy	Denoised fuzzy	[28]	Denoised [28]	[18]	[15]
Average	0.2653	0.1679	0.3377	0.1816	0.1945	0.0268
Minimum	0.2149	0.0979	0.2204	0.1031	0.1363	0
Maximum	0.3498	0.2346	0.5689	0.2768	0.3032	0.0846

Fig. 10 Examples of testing set (and its corresponding disparity maps) compared with different detection methods

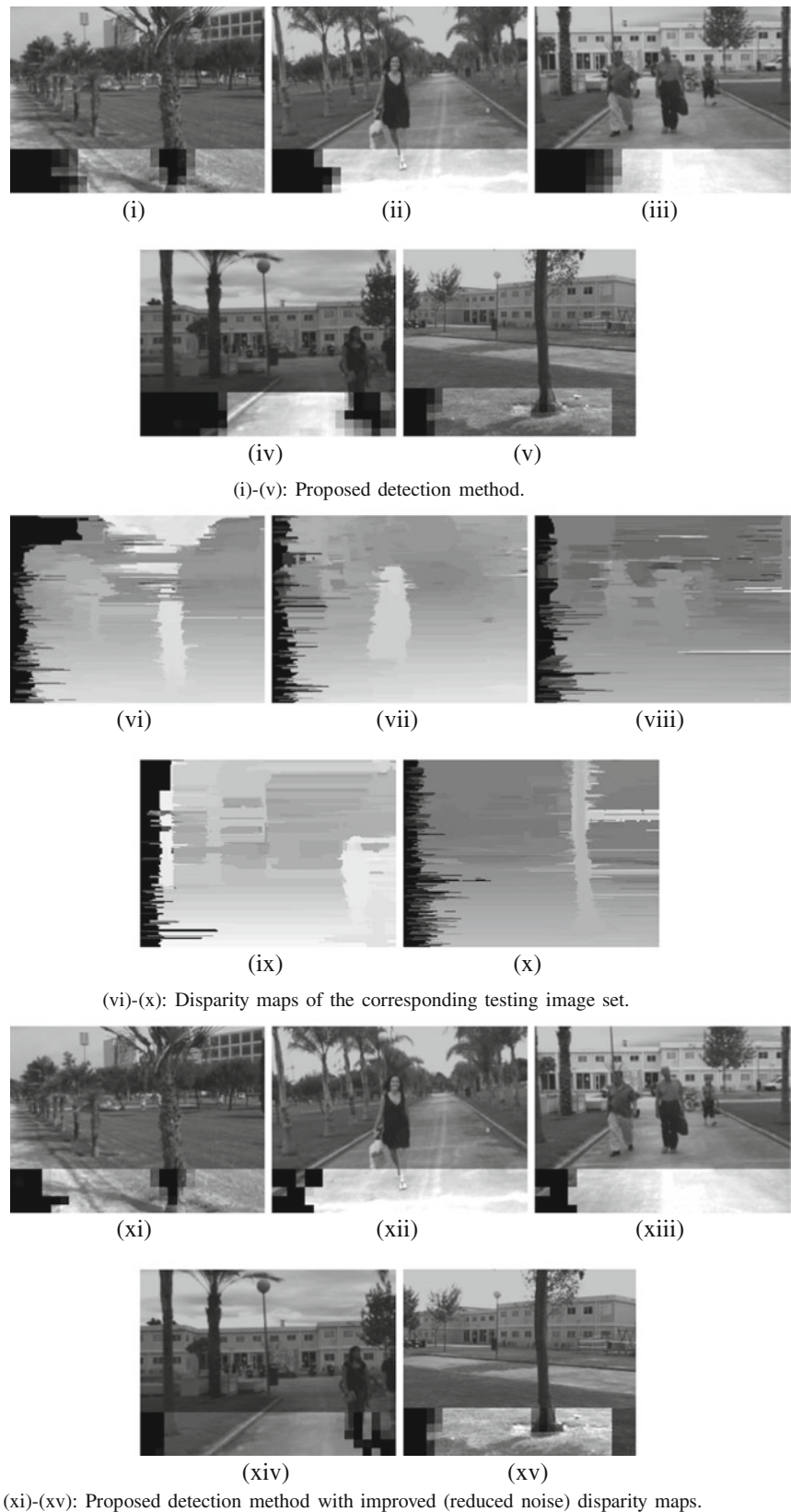
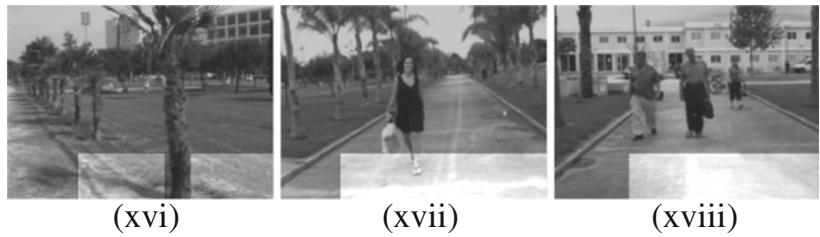


Fig. 10 (continued)

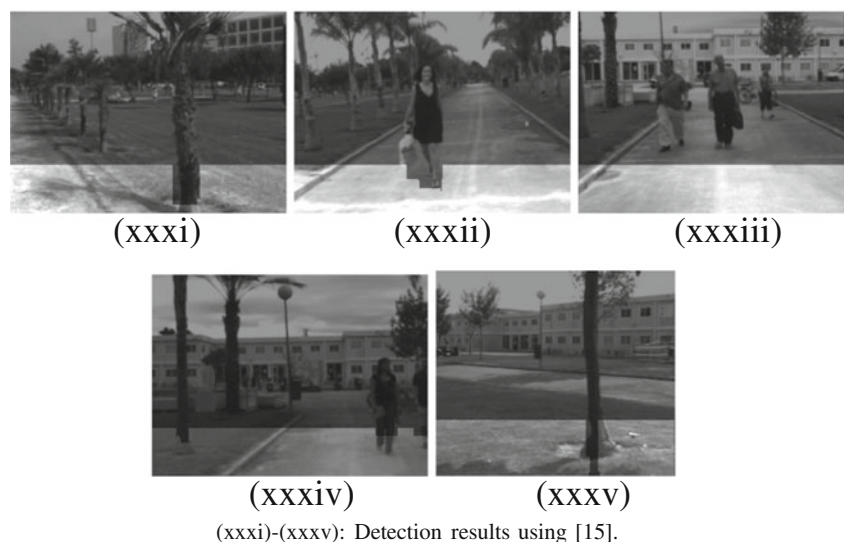
(xvi)-(xx): Crisp method proposed in [28].



(xxi)-(xxv): Crisp method proposed in [28] with improved (reduced noise) disparity maps.



(xxvi)-(xxx): Detection results using [18].

Fig. 10 (continued)

that we have abrupt depth variations when there exist depth changes greater than 5 meters. That is, by using Eq. 1, disparity variations of around 15 units from one row to another next to it. Consequently, we remove these pixels from our study before performing the planar least-squares fitting.

We can observe that the Mean Square Error decreases 46.22 % and 36.71 % for the crisp linear reference [28] and the fuzzy planar detection respectively, after denoising the disparity maps. Table 2 shows in detail the performance of the proposed method for the testing image set. Results are itemized differentiating favourable scenarios with good illumination conditions and less favourable scenarios, which have the presence of shadows and no textured areas.

In addition, Fig. 10 and Table 1 show the performances for the testing image set compared with three references to aid navigation. First and third groups of images of Fig. 10 show detection results of the proposed method without and with noise reduction in disparity maps, respectively. In the rest of the sub-figures we can compare detection results of other recent references. Comparison results of this three recent references show that the proposed method is

able to outperform most of them. The only method that presents numerically better results than the proposed in this paper is [15]. Nevertheless, this method is accurate for medium-large distances but, at very close distances, its detection results can represent a hazardous situation for the blind user (Fig. 10 (xxxv), part of the tree is detected as free space). This can be observed in Table 3, where [15] presents a larger false positives rate than the proposed method. So, [15] performs well for automotive applications in larger distances, but it is not appropriate for our system.

Moreover, the presented fuzzy algorithm is also suitable for real-time working. It has been implemented in C++ language and, under a 1.73 GHz and 1 GB RAM laptop, it spends 80 ms to obtain the dense disparity map and around 50 ms to perform the obstacle-free space detection. So, it works around 8 frames per second, which is enough for real-time performance (that is essential to be integrated as a support to navigate for blind and visually impaired people). Finally, it is also of paramount importance to remark that the whole process runs under a small laptop, without any additional GPU or FPGA (since the use of any specific hardware is not allowed due to the fact that

Table 2 Mean square error of the fuzzy detection method for the testing image set, showing detection results for favourable and unfavourable scenarios

	Favourable	Less favourable	Average
Average	0.1335	0.2874	0.1679
Minimum	0.0755	0.1222	0.0979
Maximum	0.1782	0.3451	0.2346

Table 3 False Positives and False Negatives rate for the presented method and three references to aid navigation for the testing image set

	Presented method	[28]	[15]	[18]
False Positives rate	0.0178	0.0209	0.0442	0.0203
False Negatives rate	0.1766	0.1836	0.009	0.1928

the prototype must be as lighter as possible to easy its portability).

6 Conclusions and Future Work

In this paper, a new fuzzy method for free path detection in a navigation system for visually impaired users has been proposed. The detection method is based on processing the information captured by a pair of stereo cameras and the disparity maps obtained. Disparity values vary linearly along disparity map columns in obstacle-free areas, whereas they remain nearly constant along rows in those areas. Thus, a fuzzy logic system assigns a certainty of being part of an obstacle-free path by fitting groups of disparities to a planar model. Experimental results on real outdoor images show that the method provides the most reliable results offering real-time performance.

Future work of our detection algorithm will focus on including temporal coherence in video analysis in order to perform a continuous detection. Notice that, although providing membership degrees to obstacle-free areas is more robust than results obtained from a crisp method, performing a continuous detection in conjunction with fuzzy logic will also improve results stability. Thus, we could provide the localization of the free path increasing accuracy even when very difficult environment conditions (not only in terms of lighting conditions -such as when the sun directly faces the cameras in a frame- and scene ambiguities, but also in terms of motion), doing the best to minimize the additional computational cost. Moreover, we are also planning to substitute the small laptop by a microPC to reduce the volume of the prototype and improve performances, for example, by adding texture analysis to offer more information to the end user of the system. In fact, there must exist a balance between the frame rate increment and the increasing number of applications, since the system must work at least with a minimum required frame-rate, but we can not supply too much information to blind users to avoid they get

stunned (they need more training to use the system as long as number of applications increase).

Acknowledgments N. Ortigosa acknowledges the support of Universidad Politécnica de Valencia under grant FPI-UPV 2008 and Spanish Ministry of Science and Innovation under grant MTM2010-15200. S. Morillas acknowledges the support of Universidad Politécnica de Valencia under grant PAID-05-12-SP20120696.

References

1. Cai, L., He, L., Xu, Y., Zhao, Y., Yang, X.: Multi-object detection and tracking by stereovision. *Pattern Recognit.* **43**(12), 4028–4041 (2010)
2. Hikosaka, N., Watanabe, K., Umeda, K.: Obstacle detection of a humanoid on a plane using a relative disparity map obtained by a small range image sensor. In: *Proceedings of the IEEE International Conference on Robotics and Automation*, vol. 1, pp. 3048–3053 (2007)
3. Benenson, R., Mathias, M., Timofte, R., Van Gool, L.: Fast stixel computation for fast pedestrian detection. In: *ECCV, CVVT workshop*, October (2012)
4. Huang, Y., Fu, S., Thompson, C.: Stereovision-based object segmentation for automotive applications. *EURASIP J. Appl. Signal Process.* **2005**(14), 2322–2329 (2005)
5. Duan, B.B., Liu, W., Fu, P.Y., Yang, C.Y., Wen, X.Z., Yuan, H.: Real-time on-road vehicle and motorcycle detection using a single camera. In: *IEEE International Conference on Industrial Technology*, pp. 579–584. IEEE (2009)
6. Oliveira L., Nunes, U.: On integration of features and classifiers for robust vehicle detection. In: *IEEE International Conference on Intelligent Transportation Systems*, pp. 414–419. IEEE (2008)
7. Sun, Z., Bebis, G., Miller, R.: On-road vehicle detection: A review. *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(5), 694–711 (2006)
8. Sun, H.J., Yang, J.Y.: Obstacle detection for mobile vehicle using neural network and fuzzy logic. *Neural Netw. Distrib. Process.* **4555**(1), 99–104 (2001)
9. Hui, N.B., Pratihari, D.K.: Soft computing-based navigation schemes for a real wheeled robot moving among static obstacles. *J. Intell. Robot. Syst.* **51**(3), 333–368 (2008)
10. Menon, A., Akmalawati, R., Demidenko, S.: Towards a simple mobile robot with obstacle avoidance and target seeking capabilities using fuzzy logic. In: *Proceedings IEEE Instrumentation and Measurement Technology Conference*, vol. 1–5, pp. 1003–1008 (2008)

11. Moreno-Garcia, J., Rodriguez-Benitez, L., Fernandez-Caballero, A., Lopez, M.T.: Video sequence motion tracking by fuzzification techniques. *Appl. Soft Comput.* **10**(1), 318–331 (2010)
12. Nguyen, T.H., Nguyen, J.S., Pham, D.M., Nguyen, H.T.: Real-time obstacle detection for an autonomous wheelchair using stereoscopic cameras. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2007**(1), 4775–4778 (2007)
13. Nguyen, J.S., Nguyen, T.H., Nguyen, H.T.: Semi-autonomous wheelchair system using stereoscopic cameras. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 1–20, pp. 5068–5071 (2009)
14. Grosso, E., Tistarelli, M.: Active/dynamic stereo vision. *IEEE Trans. Pattern Anal. Mach. Intell.* **17**(9), 868–879 (1995)
15. Kubota, S., Nakano, T., Okamoto, Y.: A global optimization for real-time on-board stereo obstacle detection systems. In: IEEE Intelligent Vehicles Symposium, pp. 7–12. IEEE (2007)
16. Ortigosa, N., Morillas, S., Peris-Fajarnés, G., Dunai, L.: Fuzzy free path detection based on dense disparity maps obtained from stereo cameras. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* **20**(2), 245–259 (2012)
17. Murray, D., Little, J.J.: Using real-time stereo vision for mobile robot navigation. *Auton. Robot.* **8**(2), 161–171 (2000)
18. Badino, H., Mester, R., Vaudrey, T., Franke, U.: Stereo-based free space computation in complex traffic scenarios. In: IEEE Southwest Symposium on Image Analysis & Interpretation, pp. 189–192 (2008)
19. Hoilund, C., Moeslund, T.B., Madsen, C.L., Trivedi, M.M.: Free space computation from stochastic occupancy grids based on iconic kalman filtered disparity maps. In: Proceedings International Conference on Computer Vision Theory and Applications, vol. 1, pp. 164–167 (2010)
20. Franke, U., Joos, A.: Real-time stereo vision for urban traffic scene understanding. In: IEEE Intelligent Vehicles Symposium, pp. 273–278. IEEE (2000)
21. Wedel, A., Badino, H., Rabe, C., Loose, H., Franke, U., Cremers, D.: B-spline modeling of road surfaces with an application to free-space estimation. *IEEE Trans. Intell. Transp. Syst.* **10**(4), 572–583 (2009)
22. Vergauwen, M., Pollefeys, M., Van Gool, L.: A stereo-vision system for support of planetary surface exploration. *Mach. Vis. Appl.* **14**(1), 5–14 (2003)
23. Tarel, J.P., Leng, S.S., Charbonnier, P.: Accurate and robust image alignment for road profile reconstruction. In: IEEE International Conference on Image Processing, pp. 365–368. IEEE (2007)
24. Kostavelis, I., Gasteratos, A.: Stereovision-based algorithm for obstacle avoidance. In: Lecture Notes in Computer Science, pp. 195–204. Intelligent Robotics and Applications (2009)
25. Cerri, P., Grisleri, P.: Free space detection on highways using time correlation between stabilized sub-pixel precision ipm images. In: IEEE International Conference on Robotics and Automation, pp. 2223–2228. IEEE (2005)
26. Labayrade, R., Aubert, D., Tarel, J.P.: Real time obstacle detection in stereo vision on non-flat road geometry through v-disparity representation. In: IEEE Intelligent Vehicle Symposium, pp. 646–651. INRIA (2002)
27. Ortigosa, N., Morillas, S., Peris-Fajarnés, G., Dunai, L.: Disparity maps for free path detection. In: Proceedings International Conference on Computer Vision Theory and Applications, vol. 1, pp. 310–315 (2010)
28. Ortigosa, N., Morillas, S., Peris-Fajarnés, G.: Obstacle-free pathway detection by means of depth maps. *J. Intell. Robot. Syst.* **63**(1), 115–129 (2011)
29. <http://www.casblip.com>
30. Bach y Rita, P., Collins, C., Sauters, B., White, B., Scadden, L.: Vision substitution by tactile image projection. *Nature* **221**, 963964 (1969)
31. Sampaio, E., Maris, S., Bach y Rita, P.: Brain plasticity: visual acuity of blind persons via the tongue. *Brain Res.* **908**, 204207 (2001)
32. <http://www.seeingwithsound.com>
33. Capelle, C., Trullemans, C., Arno, P., Veraart, C.: A real-time experimental prototype for enhancement of vision rehabilitation using auditory substitution. *IEEE Trans. Biomed. Eng.* **45**, 12791293 (1998)
34. Lee, S.W., Kang, S.K., Lee, S.A.: A walking guidance system for the visually impaired. *Int. J. Pattern Recognit.* **22**, 11711186 (2008)
35. Chen, C.L., Liao, Y.F., Tai, C.L.: Image-to-midi mapping based on dynamic fuzzy color segmentation for visually impaired people. *Pattern Recognit. Lett.* **32**, 549–560 (2011)
36. Lombardi, P., Zanin, M., Messelodi, S.: Unified stereo-vision for ground, road, and obstacle detection. In: Proceedings on the Intelligent Vehicles Symposium, 2005, pp. 783–788. IEEE (2005)
37. Yu, Q., Araujo, H., Wang, H.: Stereo-vision based real time obstacle detection for urban environments. In: Proceedings on the International Conference of Advanced Robotics, vol. 1, pp. 1671–1676 (2003)
38. Benenson, R., Timofte, R., Van Gool, L.: Stixels estimation without depth map computation. In: ICCV, CVVT workshop (2011)
39. Li, X., Yao, X., Murphey, Y.L., Karlsen, R., Gerhart, G.: A real-time vehicle detection and tracking system in outdoor traffic scenes. In: Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on, vol. 2, pp. 761–764 (2004)
40. Zhang, Z.Y.: A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(11), 1330–1334 (2000)
41. Dhond, U.R., Aggarwal, J.K.: Structure from stereo: a review. *IEEE Trans. Syst. Man Cybern.* **19**, 1489–1510 (1989)
42. Scharstein, D., Szeliski, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Comput. Vis.* **47**(1/2/3), 7–42 (2002)
43. Middlebury Stereo Vision Page. <http://vision.middlebury.edu/stereo/>
44. Birchfield, S., Tomasi, C.: Depth discontinuities by pixel-to-pixel stereo. *Int. J. Comput. Vis.* **17**(3), 269–293 (1999)

45. Lawrence Zitnick, C., Bing Kang, S.: Stereo for image-based rendering using image over-segmentation. *Int. J. Comput. Vis.* **75**(1), 49–65 (2007)
46. Felzenszwalb, P.F., Huttenlocher, D.P.: Efficient belief propagation for early vision. *Int. J. Comput. Vis.* **70**(1), 41–54 (2006)
47. Yang, Q., Wang, L., Yang, R., Stewnius, H., Nistr, D.: Stereo matching with color-weighted correlation, hierarchical belief propagation, and occlusion handling. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(3), 492–504 (2009)
48. Gehrig, S., Eberli, F., Meyer, T.: A real-time low-power stereo vision engine using semi-global matching. *Lect. Notes Comput. Sci.* **5815/2009**, 134–143 (2009)
49. Wedel, A., Brox, T., Vaudrey, T., Rabe, C., Franke, U., Cremers, D.: Stereoscopic scene flow computation for 3d motion understanding. *Int. J. Comput. Vis.* **95**, 29–51 (2011)
50. Hirschmuller, H.: Stereo processing by semiglobal matching and mutual information. *IEEE Trans. Pattern Anal. Mach. Intell.* **30**(2), 328–341 (2008)
51. Leung, C., Appleton, B., Sun, C.: Iterated dynamic programming and quadtree subregioning for fast stereo matching. *Image Vis. Comput.* **26**(10), 1371–1383 (2008)
52. Hartley, R.I., Zisserman, A.: *Multiple view geometry in computer vision*, 2nd edn. Cambridge University Press, ISBN: 0521540518 (2004)
53. Spiegel, M.R., Stephens, L.J.: *Statistics*, 4th edn. Mc Graw Hill (2008)
54. Kerre, E.E.: *Fuzzy sets and approximate reasoning*. Xian Jiaotong University Press (1998)
55. Dubois, D., Prade, H.: *Fuzzy sets and systems: theory and applications*. Academic Press, New York (1980)
56. Lee, C.C.: Fuzzy logic in control systems: Fuzzy logic controller-parts 1 and 2. *IEEE Trans. Syst. Man Cybern.* **20**(2), 404–435 (1990)
57. Fodor, J.C.: A new look at fuzzy-connectives. *Fuzzy Sets Syst.* **57**(2), 141–148 (1993)
58. Nalpantidis, L., Gasteratos, A.: Stereo vision for robotic applications in the presence of non-ideal lightning conditions. *Image Vis. Comput.* **28**(6), 940–951 (2010)