

Stereovision-Based Algorithm for Obstacle Avoidance

Lazaros Nalpantidis, Ioannis Kostavelis, and Antonios Gasteratos

Robotics and Automation Lab., Production and Management Engineering Dept.,
Democritus University of Thrace,
University Campus, Kimmeria, GR-671 00 Xanthi, Greece
{lanalpa, ik3339, agaster}@pme.duth.gr
<http://robotics.pme.duth.gr>

Abstract. This work presents a vision-based obstacle avoidance algorithm for autonomous mobile robots. It provides an efficient solution that uses a minimum of sensors and avoids, as much as possible, computationally complex processes. The only sensor required is a stereo camera. The proposed algorithm consists of two building blocks. The first one is a stereo algorithm, able to provide reliable depth maps of the scenery in frame rates suitable for a robot to move autonomously. The second building block is a decision making algorithm that analyzes the depth maps and deduces the most appropriate direction for the robot to avoid any existing obstacles. The proposed methodology has been tested on sequences of self-captured outdoor images and its results have been evaluated. The performance of the algorithm is presented and discussed.

Keywords: Stereo vision, obstacle avoidance, autonomous robot navigation.

1 Introduction

In this work a vision-based obstacle avoidance algorithm is presented. It is intended to be used in autonomous mobile robotics. However, the development of an efficient, solely vision-based method for mobile robot navigation is still an active research topic. Towards this direction, the first step is to avoid any obstacles through vision. However, systems placed on robots have to conform to the restrictions imposed by them. Autonomous robot navigation requires almost real-time frame rates from the responsible algorithms. Furthermore, computing resources are strictly limited onboard a robot. Thus, the omission of popular obstacle detection techniques such as the v-disparity, which require Hough-transformations, would be highly appreciated. Instead, simple and efficient solutions are demanded.

In order to achieve reliable obstacle avoidance, many popular methods involve the use of an artificial stereo vision system, due to its biomimetic nature. Stereoscopic vision can be used in order to obtain the 3D position of the depicted objects from two simultaneously captured, slightly misplaced views of a scene. Mobile robots can take advantage of stereo vision systems as a reliable method to extract information about their environment [1]. Although stereo vision provides

an enormous amount of information, most of the mobile robot navigation systems use complementary sensors in order to navigate safely [2]. The use of lasers, projectors and various other range finders is a commonplace. The goal of this work is to develop a real-time obstacle avoidance algorithm based only on a stereo camera, for autonomous mobile robots. The core of the presented approach can be divided into two separate and independent algorithms:

- *The stereo vision algorithm.* It retrieves information about the environment from a stereo camera and produces a depth image, i.e. disparity map, of the scenery.
- *The decision making algorithm.* It analyzes the data of the previous algorithm and decides the best direction, i.e. forward, right or left, for the robot to move in order to avoid any existing obstacles.

Both the algorithms have been modularly implemented in C++. The modularity of the system allows the easy modification, easy debugging and ensures the adaptability of the overall algorithm. Figure 1 presents the flow chart of the implemented algorithm. Besides the decision made, the implemented system stores input images and calculated disparity maps for possible later offline use.

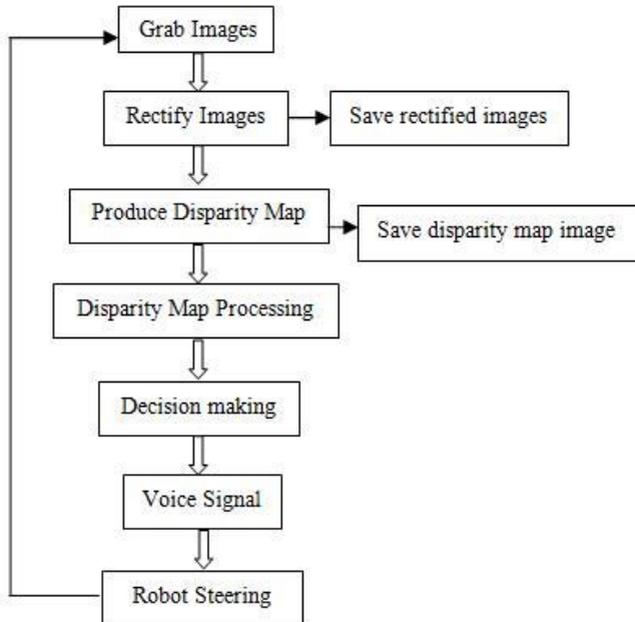


Fig. 1. Flow chart of the implemented obstacle avoidance algorithm

2 Related Work

Autonomous robots' behavior greatly depends on the accuracy of their decision making algorithms. In the case of stereo vision-based navigation, the accuracy

and the refresh rate of the computed disparity maps are the cornerstone of its success [3]. Dense local stereo correspondence methods calculate depth for almost every pixel of the scenery, talking into consideration only a small neighborhood of pixels each time [4]. On the other hand, global methods are significantly more accurate but at the same time more computationally demanding, as they account for the whole image [5]. However, since the most urgent constraint in autonomous robotics is the real-time operation, such applications usually utilize local algorithms. Muhlmann et al. in [6] describe a local method that uses the sum of absolute differences (SAD) correlation measure for RGB color images. Applying a left to right consistency check, the uniqueness constraint and a median filter, it can achieve 20 fps. Another fast Slocal AD based algorithm is presented in [7]. It is based on the uniqueness constraint and rejects previous matches as soon as better ones are detected. It achieves 39.59 fps. The algorithm reported in [8] achieves almost real-time performance. It is once more based on SAD but the correlation window size is adaptively chosen for each region of the picture. Apart from that, a left to right consistency check and a median filter are utilized.

Another possibility in order to obtain accurate results in real-time is to utilize programmable graphic processing units (GPU). In [9] a hierarchical disparity estimation algorithm is presented. On the other hand, an interesting but very computationally demanding local method is presented in [10]. It uses varying weights for the pixels in a given support window, based on their color similarity and geometric proximity. However, the execution speed of the algorithm is far from being real-time. A detailed taxonomy and presentation of dense stereo correspondence algorithms can be found in [4]. Additionally, the recent advances in the field as well as the aspect of hardware implementable stereo algorithms are covered in [11].

In the relevant literature, a wide range of sensors and various methods have been proposed in order to detect and avoid obstacles. Some interesting details about the developed sensor systems and proposed detection and avoidance algorithms can be found in [12] and [13]. The obstacle avoidance sensor systems found in literature can generally be divided into two main categories. The first category involves the use of ultrasonic sensors. They are simple to implement and can detect obstacles reliably. On the other hand, the second category involves vision-based sensor systems. This category can be further divided into the stereo vision systems (which is applied to the detection of objects in 3D) and the laser range sensors (which can be applied to the detection of obstacles both in 2D and 3D, but can barely used for real-time detection [14]). As far as the stereo vision systems are concerned, one of the most popular methods for obstacle avoidance is the estimation of the so called v -disparity image. This method requires plenty of complex calculations and is applied in order to confront the noise in low quality disparity images [15,16,17]. However, if detailed and noise-free disparity maps were available, less complicated methods could have been used instead. Considering the above as a background, the contribution of this work is the development of an algorithm for obstacle avoidance with the sole use of a stereoscopic camera. The use of only one sensor and specially

of a stereoscopic camera diminish the complexity of our system and can also be easily integrated and interact with with other stereo vision tasks such as object recognition and tracking.

3 Stereo Vision Algorithm

Contrary to most of the stereo algorithms, which directly use the camera's images, the proposed stereo algorithm uses an enhanced version of the captured images as input. The initially captured images are processed in order to extract the edges in the depicted scene. The utilized edge detecting method is the Laplacian of Gaussian (LoG), using a zero threshold. This choice produces the maximum possible edges. The LoG edge detection method smoothens the initial images with a Gaussian filter in order to suppress any possible noise. Then a Laplacian kernel is applied that marks regions of significant intensity change. Actually, the combined LoG filter, with standard deviation equal to 2, is applied at once and the zero crossings are found. The extracted edges are, afterwards, superimposed to the initial images. The steps of the aforementioned process are shown in Fig. 2. The outcome of this procedure is a new version of the original images having more striking features and textured surfaces, which facilitate the following stereo matching procedure.

The depth maps are computed using a three-stage local stereo correspondence algorithm. The utilized stereo algorithm combines low computational complexity with sophisticated data processing. Consequently, it is able to produce dense disparity maps of good quality in frame rates suitable for robotic applications. The main attribute that differentiates this algorithm from the majority of the other ones is that the matching cost aggregation step consists of a sophisticated gaussian-weighted rather than a simple summation. Furthermore, the disparity selection step is a simple winner-takes-all (WTA) choice, as the absence of any iteratively updated selection process significantly reduces the computational payload of the overall algorithm. Finally, any dedicated refinement step is also absent for speed reasons.

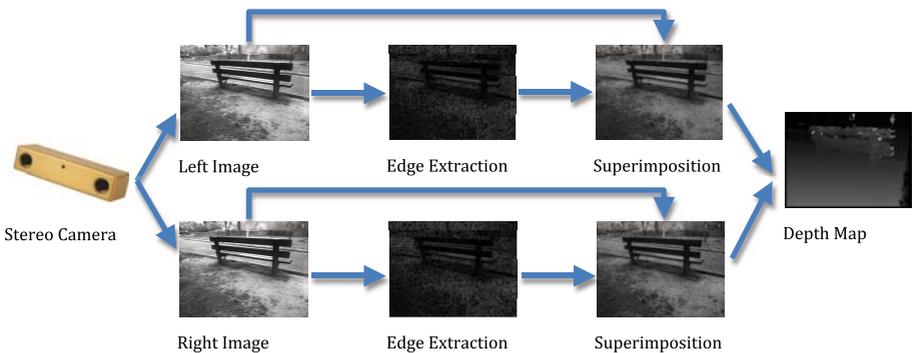


Fig. 2. Image enhancement steps of the proposed stereo algorithm

The matching cost function utilized is the truncated absolute differences (AD). AD is inherently the simplest metric of all, involving only summations and finding absolute values. The AD are truncated if they exceed the 4% of the maximum intensity value. Truncation suppresses the influence of noise in the final result. This is very important for stereo algorithms that are intended to be applied to outdoor scenes. Outdoor pairs usually suffer from noise induced by a variety of reasons, e.g. lighting differences and reflections. For every pixel of the reference (left) image, AD are calculated for each of its candidate matches in the other (right) image.

The computed matching costs for every pixel and for all its potential disparity values comprise a 3D matrix, usually called as disparity space image (DSI). The DSI values for constant disparity value are aggregated inside fix-sized square windows. The dimensions of the chosen aggregation window play an important role in the quality of the final result. Generally, small dimensions preserve details but suffer from noise. On the contrast, large dimensions may not preserve fine details but significantly suppress the noise. This behavior is highly appreciated in outdoor robotic applications where noise is a major factor, as already discussed. The aggregation windows dimensions used in the proposed algorithm are 19x19 pixels. This choice is a compromise between real-time execution speed and noise cancelation. The AD aggregation step of the proposed algorithm is a weighted summation. Each pixel is assigned a weight depending on its Euclidean distance from the central pixel. A 2D Gaussian function determines the weights value for each pixel. The center of the function coincides with the central pixel. The standard deviation is equal to the one third of the distance from the central pixel to the nearest window-border. The applied weighting function can be calculated once and then be applied to all the aggregation windows without any further change. Thus, the computational load of this procedure is kept within reasonable limits.

Finally, the optimum disparity value for each pixel, i.e. the disparity map, is chosen by a simple, non-iterative WTA step. In the resulting disparity maps, smaller values indicate more distant objects, while bigger disparity values indicate objects lying closer.

4 Decision Making Algorithm

The previously calculated disparity map is used to extract useful information about the navigation of a robot. Contrary to many implementations that involve complex calculations upon the disparity map, the proposed decision making algorithm involves only simple summations and checks. This is feasible due to the absence of significant noise in the produced disparity map. The goal of the developed algorithm is to detect any existing obstacles in front of the robot and to safely avoid it, by steering the robot left, right or to moving it forward.

In order to achieve that, the developed method divides the disparity map into three windows, as in Fig. 3.

In the central window, the pixels p whose disparity value $D(p)$ is greater than a defined threshold value T are enumerated. Then, the enumeration result is



Fig. 3. Depth map's division in three windows

examined. If it is smaller than a predefined rate r of all the central windows pixels, this means that there are no obstacles detected exactly in front of the robot and in close distance, and thus the robot can move forward. On the other hand, if this enumeration's result exceeds the predefined rate, the algorithm examines the other two windows and chooses the one with the smaller average disparity value. In this way the window with the fewer obstacles will be selected. The pseudocode of the implemented simple decision making algorithm follows:

Decision Making Pseudocode

```

for all the pixels p of the central window {
  if  $D(p) > T$  {
    counter++ }
  numC++ }
if counter <  $r\%$  of numC {
  GO STRAIGHT }
else {
  for all the pixels p of the left window {
    sumL +=  $D(p)$ 
    numL++ }
  for all the pixels p of the right window {
    sumR +=  $D(p)$ 
    numR++ } }
avL = sumL / numL
avR = sumR / numR
if avL < avR {
  GO LEFT }
else {
  GO RIGHT } }

```

The values of the parameters T and r play an important role to the algorithm's behavior. Small values of T in conjunction with small values of r favor the

hesitancy in moving forward, ensuring obstacle avoidance but at the same time being susceptible to false alarms due to noise. On the other hand, the opposite scenario is less susceptible to false alarms but may be proven risky. The values of T and r were empirically defined as a compromise between the two aforementioned behavioral patterns. The decided values were $T = 120$ and $r = 20\%$. In the example of Fig. 3 the number of pixels in the central window whose disparity value is greater than 120 is 40160 which is more than the 20% of all the window's pixels (i.e. 8160). Consequently, the robot is not allowed to move forward and has to examine the other two possibilities. The robot will decide to turn left as the left window has smaller average disparity values than the right one.

5 Experimental Validation

The performance of the examined methodology for obstacle avoidance has been examined by applying it on a sequence of 25 stereo image pairs. The experiments took place in outdoor scenes and the obstacles were natural elements like trees, bushes, benches, pavements and other objects found in the captured scenes. Various single image pairs, as well as routes were captured. That is, the algorithm's decision concerning a test pair, dictates the next pairs' capture position. A series of such linked pairs comprise a route. All the 25 test image pairs are available to be freely downloaded from [18]. Figure 4 presents a sample route comprising of 7 image pairs. For each image pair the reference image is shown on the left, the calculated disparity map in the middle, and the direction decided by the algorithm is shown on the right. The decision for each image pair dictated the camera's next position, as easily deduced by the sequence of Fig. 4.

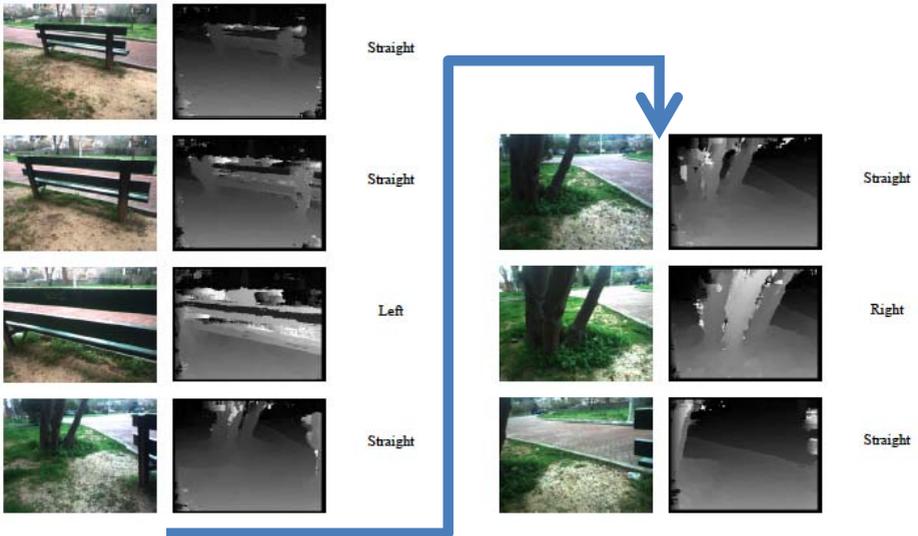


Fig. 4. A sample outdoor route and the algorithm's outputs

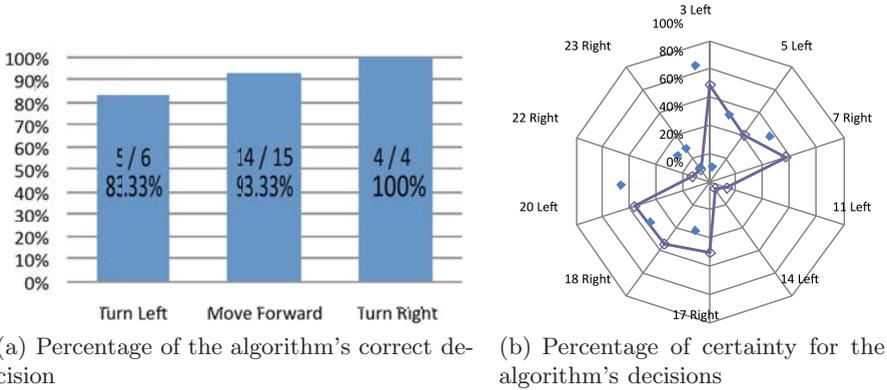


Fig. 5. Results of the proposed algorithm

The algorithm's results, when applied on the 25 test image pairs, are summarized in Fig. 5. Figure 5(a) presents the percentage of the algorithm's correct decisions for each possible direction. A correct decision is not only one that avoids collision with an existing obstacle, but at the same time chooses the most preferable direction. As shown, the algorithm is more than 83% of the times correct, for any direction. Actually the algorithm has been correct in 23 out of the 25 cases, i.e. 92% of the times.

On the other hand, the certainty of each decision is of interest as well. Due to the nature of the proposed algorithm, as presented in the previous sections, the decision about moving forward is based on some parameters' values and is totally independent from the other two possible decisions. Consequently, a measure of certainty would be meaningful only in the cases of the left or right decisions. As these two decisions are based on the same heuristic, they can be directly compared. The certainty *cert* of a direction's decision which yields an average disparity avD_1 over the other direction which yields $avD_2 > avD_1$ is calculated as:

$$cert = \frac{avD_2 - avD_1}{avD_2} \tag{1}$$

The results for the left and right decisions of the algorithm are shown in Fig. 5(b). For each decision the pair's indicating number as well as the algorithm's decision is given. The certainty ranges form 0% for no certainty at all, to 100% for absolute certainty. The bigger the area defined by the resulting points, the bigger the algorithm's overall certainty. However, big values of certainty are not always achievable. In the extreme case when both the left and the right direction are fully traversable, the certainty measure would become 0%. Despite this fact, the certainty is useful. Observing the correlation between false decisions and certainty values, a threshold could be decided, below which the algorithm should reconsider its decision.

6 Conclusion

A vision-based obstacle avoidance algorithm for autonomous mobile robots was presented. The proposed algorithm requires only one sensor, i.e. a stereo camera, and a low amount of involved computations. The algorithm's structure consists of a specially developed and optimized stereo algorithm that produces noise-free depth maps, and a computationally simple decision making algorithm. The decision making algorithm avoids complex calculations and transformations. Consider as an example, the case of the popular v-disparity implementation where Hough-transformation is needed in order to compensate for the low quality disparity maps. On the other hand, simpler than the proposed direction deciding algorithms fail to yield correct results. In this case, consider an algorithm where the three windows of Fig. 3 are treated equally and the smallest average disparity is sought. This methodology is doomed to fail in the case, among many others, where only a thin obstacle is close to the robot and other obstacles are in medium range. Such a naive algorithm would chose the direction towards the close thin obstacle, avoiding the medium ranged obstacles.

The proposed methodology has been tested on sequences of self-captured outdoor images and its results have been evaluated. Its performance has been presented and discussed. The proposed algorithm managed to avoid the obstacles successfully in the vast majority of the tested image pairs. Despite its simple calculations, both during the disparity map generation and the decision making, the algorithm exhibited promising behavior. The simple structure and the absence of heavy computational payload are characteristics highly desirable in autonomous robotics. The real-time collision-free navigation of autonomous robotic platforms is the first step towards the accomplishment of more complex activities, e.g. path planning and mapping of an area. Consequently, the proposed algorithm is suitable for autonomous robotic applications and is able to provide real-time obstacle avoidance behavior, based solely on a stereo camera.

Acknowledgments. This work was supported by the E.C. funded research project for vision and chemiresistor equipped web-connected finding robots, "View-Finder", FP6-IST-2005-045541.

References

1. Iocchi, L., Konolige, K.: A multiresolution stereo vision system for mobile robots. In: Italian AI Association Workshop on New Trends in Robotics Research (1998)
2. Siegwart, R., Nourbakhsh, I.R.: Introduction to Autonomous Mobile Robots. MIT Press, Massachusetts (2004)
3. Schreer, O.: Stereo vision-based navigation in unknown indoor environment. In: 5th European Conference on Computer Vision, vol. 1, pp. 203–217 (1998)
4. Scharstein, D., Szeliski, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision* 47(1-3), 7–42 (2002)
5. Torra, P.H.S., Criminisi, A.: Dense stereo using pivoted dynamic programming. *Image and Vision Computing* 22(10), 795–806 (2004)

6. Muhlmann, K., Maier, D., Hesser, J., Manner, R.: Calculating dense disparity maps from color stereo images, an efficient implementation. *International Journal of Computer Vision* 47(1-3), 79–88 (2002)
7. Di Stefano, L., Marchionni, M., Mattoccia, S.: A fast area-based stereo matching algorithm. *Image and Vision Computing* 22(12), 983–1005 (2004)
8. Yoon, S., Park, S.K., Kang, S., Kwak, Y.K.: Fast correlation-based stereo matching with the reduction of systematic errors. *Pattern Recognition Letters* 26(14), 2221–2231 (2005)
9. Zach, C., Karner, K., Bischof, H.: Hierarchical disparity estimation with programmable 3d hardware. In: *International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*, pp. 275–282 (2004)
10. Yoon, K.J., Kweon, I.S.: Adaptive support-weight approach for correspondence search. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28(4), 650–656 (2006)
11. Nalpantidis, L., Sirakoulis, G.C., Gasteratos, A.: Review of stereo vision algorithms: from software to hardware. *International Journal of Optomechatronics* 2(4), 435–462 (2008)
12. Borenstein, J., Koren, Y.: Real-time obstacle avoidance for fast mobile robots in cluttered environments. *IEEE Transactions on Systems, Man, and Cybernetics* 19(5), 1179–1187 (1990)
13. Ohya, A., Kosaka, A., Kak, A.: Vision-based navigation of mobile robot with obstacle avoidance by single camera vision and ultrasonic sensing. *IEEE Transactions on Robotics and Automation* 14(6), 969–978 (1998)
14. Vandorpe, J., Van Brussel, H., Xu, H.: Exact dynamic map building for a mobile robot using geometrical primitives produced by a 2d range finder. In: *IEEE International Conference on Robotics and Automation*, Minneapolis, USA, pp. 901–908 (1996)
15. Labayrade, R., Aubert, D., Tarel, J.P.: Real time obstacle detection in stereovision on non flat road geometry through "v-disparity" representation. In: *IEEE Intelligent Vehicle Symposium*, Versailles, France, vol. 2, pp. 646–651 (2002)
16. Zhao, J., Katupitiya, J., Ward, J.: Global correlation based ground plane estimation using v-disparity image. In: *IEEE International Conference on Robotics and Automation*, Rome, Italy, pp. 529–534 (2007)
17. Soquet, N., Aubert, D., Hautiere, N.: Road segmentation supervised by an extended v-disparity algorithm for autonomous navigation. In: *IEEE Intelligent Vehicles Symposium*, Istanbul, Turkey, pp. 160–165 (2007)
18. Nalpantidis, L., Kostavelis, I.: *Group of Robotics and Cognitive Systems* (2009), <http://robotics.pme.duth.gr/reposit/stereoroutes.zip>