Integrating the whole cost-curve of stereo into occupancy grids

Martim Brandão, Ricardo Ferreira, Kenji Hashimoto, José Santos-Victor and Atsuo Takanishi

Abstract— Extensive literature has been written on occupancy grid mapping for different sensors. When stereo vision is applied to the occupancy grid framework it is common, however, to use sensor models that were originally conceived for other sensors such as sonar. Although sonar provides a distance to the nearest obstacle for several directions, stereo has confidence measures available for each distance along each direction. The common approach is to take the highest-confidence distance as the correct one, but such an approach disregards mismatch errors inherent to stereo.

In this work, stereo confidence measures of the whole sensed space are explicitly integrated into 3D grids using a new occupancy grid formulation. Confidence measures themselves are used to model uncertainty and their parameters are computed automatically in a maximum likelihood approach. The proposed methodology was evaluated in both simulation and a real-world outdoor dataset which is publicly available. Mapping performance of our approach was compared with a traditional approach and shown to achieve less errors in the reconstruction.

I. INTRODUCTION

Occupancy grids [1] provide an excellent tool for world mapping from sensor measurements, which is particularly useful for robot navigation and motion planning. The concept was initially proposed for use with sonar sensors and later on also applied to stereo vision. However, attempts to integrate stereo into occupancy grids have mainly opted for the same sensor models as the ones used for sonar.

In stereo vision, 3D information (i.e. object positions in the world) is extracted from 2D measurements (i.e. images from two cameras) by a process called stereo matching. For each pixel on one image, a line on the other image is scanned for the matching pixel. Each pixel along this line then corresponds to a 3D point in the world. Cost functions are used to assign a confidence to each match hypothesis and this vector of costs along one line is usually called the cost-curve. The existence of a confidence measure for each distance means more information is available. However, attempts to integrate stereo into occupancy grids have opted

to take the least-cost match (i.e. minimum of the cost-curve) as the correct one in a winner-take-all approach [2], [3], [4], [5], [6]. Unfortunately, the inherent challenges in stereo make it difficult to guarantee the effectiveness of a winner-take-all approach. In order to solve this problem, filtering of stereo results is done by discarding matches where, for example, the minimum cost "peak" is not salient enough or texture is low. Then a unimodal probability distribution is usually centered around these least-cost points to model their uncertainty. However, because unimodal distributions are used, the stereo mismatch problem is not dealt with.

Using confidence measures of stereo matching to model uncertainty, and hence occupancy probability, can pave way for better mapping. Not only that, if the whole stereo cost-curve contributes to occupancy of cells, then the correspondence (mismatch) problem is implicitly dealt with. Possible correct matches that do not correspond to global minima due to noise still contribute to grid occupancy; while wrong minimum costs are filtered out due to inconsistency. It is with this in mind that we conducted the present work.

Our contribution is as follows:

- A new occupancy grid formulation is proposed to better deal with sensors modeled by multimodal probability distributions such as stereo, integrating not only the least-cost estimate but the whole cost-curve.
- Parameters of the stereo confidence function used to model stereo uncertainty are computed automatically in a maximum likelihood approach.
- The proposed methodology was evaluated in both simulation and real-world outdoors datasets [7]. Mapping performance was compared with a traditional, winner-take-all approach, and shown to outperform it: both empirically and in terms of reconstruction error.

The paper introduces related work in Section I-A. Background in occupancy grid mapping is given in II and the proposed approach described in III. We finish with an experimental evaluation and conclusions.

A. RELATED WORK

Occupancy grids were initially proposed in [8]. Through this framework, a world map is built given sensor measurements, sensor position in the world and a sensor model. The map is defined as a grid of cells which can be in an occupied or free state. This is done with a probabilistic approach, accounting for uncertainty in the sensors. Occupancy grids have been used successfully for robot navigation using various sensors such as sonar [1], [9], laser rangefinders and also stereo vision [2], [3], [4], [5], [6].

When integrating stereo into occupancy grids, uncertainty of the sensor has been modeled in different ways. Works such as [10] assume a constant likelihood of occupation at the least cost stereo matches. A fixed "occupied contribution" constant is added to cells with projected least-cost matches; and another constant is subtracted from cells on the ray from these matches to the camera. The constants are chosen empirically or by observation. In [2] a sensor model for stereo is defined, accounting for the increase of uncertainty with distance from the camera and also with distance to edges in the image (refered to as "inference errors" of assuming that line segments correspond to surfaces). Correspondence errors are disregarded. Kalman filters are used in [6] to track stereo matches and the filters' covariance matrix is used to estimate the uncertainty of points and compute occupancy probability. An inverse sensor model of stereo for occupancy grids is proposed in [4] which accounts for the increase of uncertainty with distance from the sensor. The resulting occupancy profile along a ray according to this model is very similar to that of [1]. Another attempt to model the uncertainty associated with stereo is done in [5]. The correspondence problem is dealt with but only by considering probability constants representing the probability of false positives and false negatives on a stereo measurement.

According to [11], disparity estimates have gaussian distributed error with small standard deviation below 1 pixel. However, as pointed by [3], this study disregards stereo mismatches. Correspondence mismatches are usually filtered out by a left-right consistency check or the cost-curve's peak-ratio (if the cost minimum is not "salient" enough). In [3] the problem is tackled by filtering out disparity "spikes" using image filters.

However, in these approaches it has been common practice to model the stereo sensor in the same way as a laser and sonar: that is, as a sensor returning a single distance whose uncertainty is a unimodal probability distribution. Only the least-cost stereo match is considered, and the rest of the cost-curve is not used when dealing with the mismatch problem. Thorough reviews on stereo matching confidence measures and their effectiveness on modeling true disparity have been published recently [12], which also motivates our research.

II. OCCUPANCY GRID MAPPING OVERVIEW

Occupancy grid mapping consists of building a map of the world from sensor measurements and sensor position in that world. It can be defined as finding the probability $p(m|z_{1...t}, x_{1...t})$ where m is the map, z the measurements, x sensor position and t instant of time. Each map cell m_{XYZ} can have one of two states, occupied O or free $\overline{O}$.

Measurements from one or more sensors can be fused and integrated in time iteratively from the inverse sensor model $p(O|z)$ and an occupancy prior $p(O)$ (how likely it is for a cell to be occupied).

A. Elfes' original formulation of the inverse sensor model

Consider the 1-dimensional grid G of N cells obtained by intersection of a sensor ray with the map m. The range

sensor measurement is a cell index $r \in \{1, ..., N\}$ of the first obstacle along the ray.

As proposed by Elfes [8], occupancy of cell $i \in \{1, ..., N\}$ is computed from range measurement r as

$$p(Oi|r) = \frac{p(r|O_i)p(O_i)}{p(r|O_i)p(O_i) + p(r|\overline{O}_i)p(\overline{O}_i)}, \qquad (1)$$

where O_i represents the event of cell i being occupied.

Fix cell i as occupied and consider all possible configurations of the grid G by varying all the other cell's states. Let configuration number k be represented by $Conf_k(G, O_i)$. There are two possible states for each cell and so 2^{N-1} configurations exist. Since $p(r|O_i)$ depends on the state of cells other than cell i, the original formulation of occupancy grids computes it by total probability,

$$p(r|Oi) = \sum_k p(r|O_i, Conf_k(G, O_i))p(Conf_k(G, O_i)), \tag{2}$$

where the sum is made over all grid configurations. $p(r|\overline{O}_i)$ can be computed using the same process and $p(Conf_k(G, O_i))$ is set constant and equal to 0.5^{N-1}.

Take r' as the real cell index of the first obstacle. The direct sensor model $p(r|r')$ is usually modelled by a normal distribution of the error, with mean r'. What r' represents in terms of occupancy is that cell r' is occupied and all cells on the interval $\{1, ..., r' - 1\}$ are free. Hence, a range sensor's direct sensor model is usually given by:

$$p(r|O_{r'}, \overline{O}_{r'-1}...\overline{O}_1) \sim \mathcal{N}(r', var). \tag{3}$$

The result of applying (1) (2) (3) to a sensor model with $var = 5$ is displayed in Fig. 1. Other occupancy grid formulations have been proposed, including for stereo vision [4], in general assuming a similar resulting shape of this 1D occupancy profile.
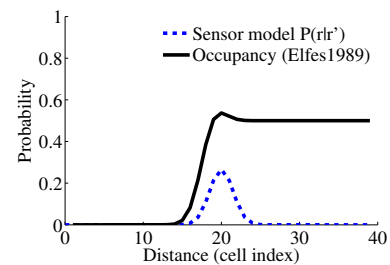

Fig. 1. Occupancy profile given Elfes' original formulation using a sensor model with normally distributed error and variance 5. The range measurement was r=20.

This formulation was made with unimodal sensor models in mind. Its direct application to multimodal probability distributions, however, leads to unwanted results. In equation (2), there will be many more configurations with a target (first occupied cell) close to the sensor than with targets further away. The contribution of cells hence increases drastically with proximity to the sensor, leading mostly to global maxima on the first (local) maximum, as will be shown in the next section.

III. OCCUPANCY GRIDS FROM STEREO USING THE WHOLE COST-CURVE

In our approach each cell's occupancy is computed from a stereo matching cost along the cost-curve. The probability distribution of occupancy is therefore multimodal. The fact that only the local maxima closesest to the sensor contribute to occupancy of cells on the previous formulation is, however, undesired. This behaviour is shown in Fig. 2. We deal with this problem by proposing the following occupancy grid formulation.

A. Occupancy grid formulation

In stereo vision the sensor measures a cost of treating each cell in $\{1, ..., N\}$ as the target (first obstacle along the ray), as opposed to measuring a cell index of the target.

Consider the cost vector $E = c_1, c_2, ..., c_N$ of assigning the target to each cell. To obtain the inverse sensor model we compute $p(O_i|E)$. Also, to avoid long equations we will represent $(\overline{O}_{i-1}...\overline{O}_2, \overline{O}_1)$, which can be seen as visibility of cell i, by V_i. By the law of total probability,

$$p(O_i|E) = p(O_i|EV_i)p(V_i|E) + p(O_i|E\overline{V}_i)(1 - p(V_i|E)). \quad (4)$$

The term $p(V_i|E)$ can be computed by recursively applying the definition of conditional probability,

$$\begin{aligned} p(V_i|E) &= p(V_{i-1}\overline{O}_{i-1}|E) = p(\overline{O}_{i-1}|V_{i-1}E)p(V_{i-1}|E) \\ &= p(\overline{O}_{i-1}|V_{i-1}E)p(\overline{O}_{i-2}|V_{i-2}E)p(V_{i-2}|E) \\ &= ... = \prod_{j=0...i-1} p(\overline{O}_j|EV_j). \end{aligned} \quad (5)$$

In the event that a cell is not visible, (stereo) costs can give no information about occupancy on that cell. By considering that the event of an occupied cell O_i and the cost curve E are conditionally independent given the cell is not visible $\overline{V}_i$, then the prior $p(O_i|E\overline{V}_i) = p(Oi|\overline{V}_i)$. This prior models our expectation on the geometry of the physical environment, in the same way $p(O_i)$ does in other models of the state of the art. We set the prior to be uniform and equal to 0.5 for equal probability of occupied and free cells.

The term $p(O_i|EV_i)$ corresponds to our inverse sensor model. Through Bayes theorem,

$$p(O_i|EV_i) = \frac{p(E|O_iV_i)p(O_i|V_i)}{p(E|O_iV_i)p(O_i|V_i) + p(E|\overline{O}_iV_i)p(\overline{O}_i|V_i)}. \quad (6)$$

Fig. 2 shows the result of applying both the original formulation and our formulation of occupancy grids to a given inverse sensor model $p(O_i|EV_i)$.

For unimodal sensor models both formulations lead to identical occupancy profiles. However, as previously pointed, the original formulation typically leads to a global maximum on the first peak when multimodal sensor models are used.

B. Stereo matching

Consider two images I_1 and I_2, aligned along the x axis. In stereo vision, the cost-curve $E(d)$ of assigning $I_2(x, y)$ to $I_1(x + d, y)$ is computed for each pixel (x, y). The (conditional) probability function of measuring $E(d)$ at the
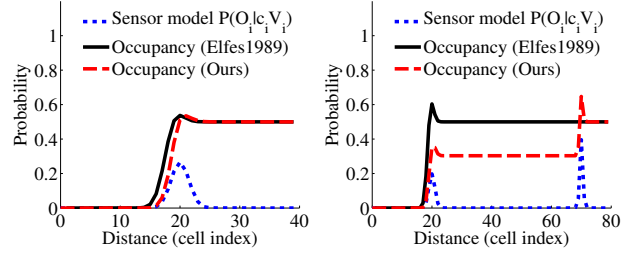

Fig. 2. Occupancy profile given the original formulation and our formulation, when both a unimodal and multimodal sensor model is used. c_i represents the cost measurement on cell i. V_i visibility of cell i.

true disparity can be defined assuming a normal distribution of costs [13],

$$p(E(d)|d) = \frac{e^{-\frac{E(d)}{2\sigma^2}}}{\sum_{i=1...n} e^{-\frac{E(i)}{2\sigma^2}}}, \quad (7)$$

where σ represents expected image pixel error. Although out of scope of this paper, different confidence measures also exist to compute $p(E(d)|d)$. For a thorough review please refer to [12]. In that review, (7) ranks within the highest confidence measures considering the whole cost-curve.

Let D refer to the cell index, in cartesian space, corresponding to disparity d. Then as a range sensor, stereo's sensor model can be written as

$$p(E(d)|d) = p(E(d)|O_D, V_D). \quad (8)$$

C. Stereo inverse sensor model

The sensor model term $p(E|O_iV_i)$ was already computed in (7). The model $p(E|\overline{O}_iV_i)$ of stereo costs at free-space (i.e. cells which are visible and free) should also be defined. An estimate of this model could be drawn from the cost-volume of all visible pixels. Fig. 3 shows this distribution taken from the example stereo pair "Cones" in the Middlebury database [14] given ground truth disparity. The figure indicates a normal distribution of costs.
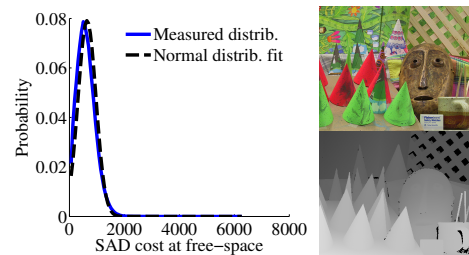

Fig. 3. Probability distribution of SAD (Sum of Absolute Differences) costs on a 7x7 window, measured from all visible pixels of the "Cones" stereo pair.

We also considered the following approximation and compare it in Fig. 4. Assuming that $p(E|V_i)$ is uniformly distributed (every cost measurement is equally probable at visible cells) and that $\sum_i p(O_i|EV_i) = 1$, then the denominator of (6) is constant and the equation reduces to

$$p(O_i|EV_i) = p(E|O_iV_i). \quad (9)$$

Although the uniform distribution approximation seems strong we found out in our experiments that it actually leads to similar occupancy profiles, thus leading to believe that a free-space model does not add considerable information to the range model.

Fig. 4 compares the output occupancy along a camera ray given: 1) the original occupancy grid formulation, 2) our formulation with free-space model estimation and 3) our formulation using approximation (9). The results show that our formulation can deal with multiple maxima, generating an occupancy profile with high occupancy near salient confidence maxima. Approximation (9) also leads to similar results when compared to the use of a free-space model.

The conditional occupancy prior $p(O_i|V_i)$ was set to 0.5 for equal probability of occupied and free cells.
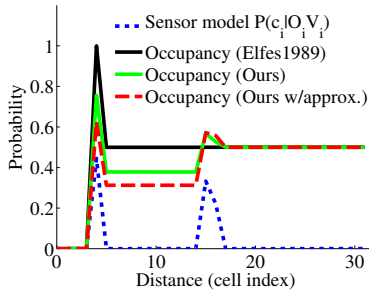

Fig. 4. Occupancy along a 1D ray given Elfes' formulation and our formulation with and without approximation (9)

D. Model parameter estimation

The σ parameter in (7) is usually estimated manually by observation. In this work, nevertheless, we opt to estimate the parameter automatically from maximum likelihood for every stereo pair, with robustness to scene and lighting conditions in mind. Given a normal distribution, the ML parameter σ^2 is given by

$$\hat{\sigma^2} = \sum (E_1 - \overline{E_1})^2, \qquad (10)$$

where $E_1 = I_2(x, y) - I_1(x + d', y)$, $\overline{E_1}$ is the average of E_1 for all pixels and d' the winner-take-all disparity. Practically, this means that one can compute the variance of the error at the same time the stereo cost-volume is being computed from the pixel difference.

IV. EXPERIMENTAL RESULTS

The proposed method was tested on both virtual and real-world outdoors image data. Our goal is to test whether the integration of the whole cost-curve of stereo matching leads to better grid reconstructions. We implemented the algorithms on 3D grids for better visibility of the improvements since less matches are accumulated into each cell.

The winner-take-all approach was implemented using the formulation presented in Section II-A. The parameter σ of our approach was computed for each stereo pair as described in Section III-D, from all global minimum stereo matches. For both approaches, a SAD (Sum of Absolute Differences) on a 7x7 window was chosen as the stereo matching cost.

Approximation (9) was used in our method for simplicity. Projection of stereo matches on the occupancy grid is made through line drawing: from sensor space to grid space. When multiple points fall on the same cell, the maximum of their occupancy probabilities is taken (obstacle priority). For low disparities distance is high and so cells between two consecutive disparities are filled assuming a uniform distribution.

A. Virtual environment

In this section we show the success of our approach when applied to scenes with low texture and repetitive patterns.

A straight-forward application of the proposed formulation is a scene with vertical repetitive characteristics. The Peak Ratio (c_{m2}/c_{m1}) of the cost curves will be low, thus leading to either false positives or holes in the reconstruction depending on the Peak Ratio threshold choosen when the winner-take-all approach is used. On the other hand, a whole-cost-curve approach is expected to keep the occupancy probability at repetitions high enough, and eliminate false-positives with time, as the viewing angle changes.

To empirically confirm this hypothesis we simulated a simple environment with thin vertical bars, camera moving around them. Fig. 5 shows the resulting reconstruction of the scenario after 20 frames of camera motion using our method. Blue regions indicate occupied cells, which should form parallel bars. On the figure, cells are drawn on top of the point cloud obtained from least-cost stereo matches.
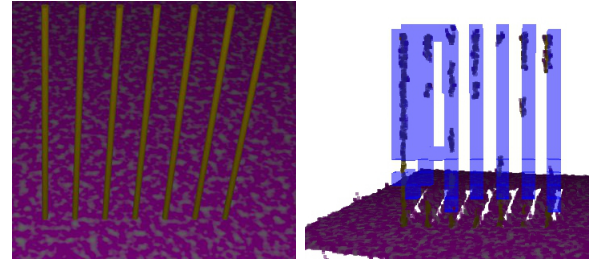

Fig. 5. Virtual scenario with vertical bars to induce similar cost minima (left); resulting occupancy grid in our whole-cost-curve approach (right). Occupied cells are marked with blue. Result should be 7 parallel bars.

Fig. 6 shows the results after the same number of frames from the winner-take-all approach using three different stereo filtering thresholds (Peak Ratio). High confidence restrictions lead to holes in the reconstructions. Less filtering however leads to more errors and intensive post-filtering is needed. A threshold leading to a reconstruction with no holes and no outliers was not found. The image in the center reveals that there are still holes in the reconstruction when outliers start appearing on a traditional winner-take-all approach.

Our proposed approach achieved full reconstruction without outliers and does not require manual calibration.

B. Real outdoors datasets

The proposed method was also evaluated on the KITTI outdoors driving dataset "2011-09-28 drive 0038" and "2011-09-28 drive 0045" [7]. The 109 and 41 frames of the
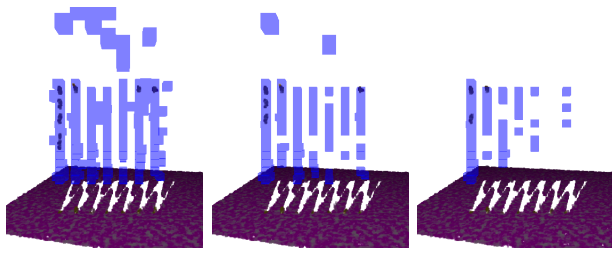
Fig. 6. Resulting occupancy grid computed in a traditional winner-take-all approach when using three different filtering thresholds (Peak Ratio): 1.4 (left) 1.5 (center) and 1.6 (right). Occupied cells are marked with blue. The result should be 7 parallel bars.

datasets were used to compare performance of occupancy grids when using whole-cost-curve and winner-take-all stereo approaches. The first dataset is shown in Fig. 7. In order to obtain a ground-truth grid, a standard occupancy grid algorithm for range data was implemented and run on all frames using the available range-finder data. After the process was run through all frames, cells with resulting occupancy odds over a threshold are considered occupied and the rest as free. The localization data, given by the dataset, was assumed to be correct. Cell size used was 20cm x 20cm x 20cm and the resulting grid 160m x 10m x 4m. Since these datasets contain moving obstacles, we decided to mark any cell which was occupied during the image sequence as occupied in the ground-truth grid, including moving obstacles. These cells were not counted as mistakes in the reconstruction. Generated ground-truth is shown on Fig. 7.
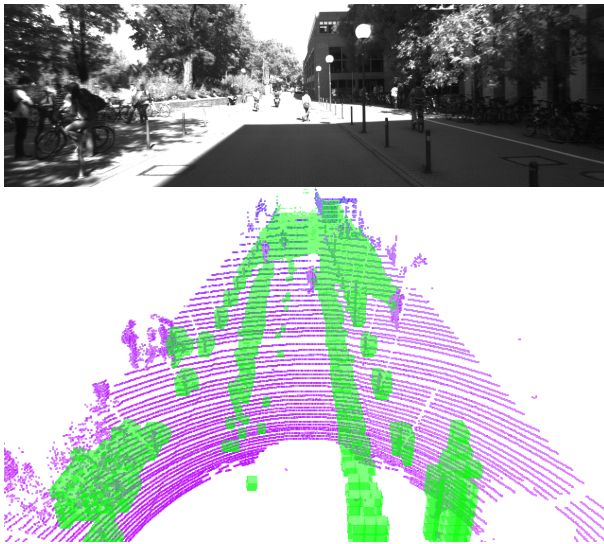

Fig. 7. KITTI dataset "2011-09-28 drive 0038" was used for evaluation of the proposed methodology on a real outdoor scenario. Initial image from the dataset (top); Laser reconstruction with overlapped occupied grid cells in green (bottom).

After running both methods on the stereo data, cells with resulting occupancy probability over 0.5 were considered occupied. Resulting occupancy grids are shown in Fig. 8. A higher number of correct matches can be observed for our proposed approach, especially on the poles on the left side of

the street. The reason for these not showing on the winner-take-all approach could be wrong distant matches (low disparity) that contribute to large areas of free space. The existance of other similar-confidence matches on the way to these points would however lead to occupancy profiles closer to P=0.5 on the whole-cost-curve approach and hence faster convergence to occupancy when the matches' confidence measure becomes higher.
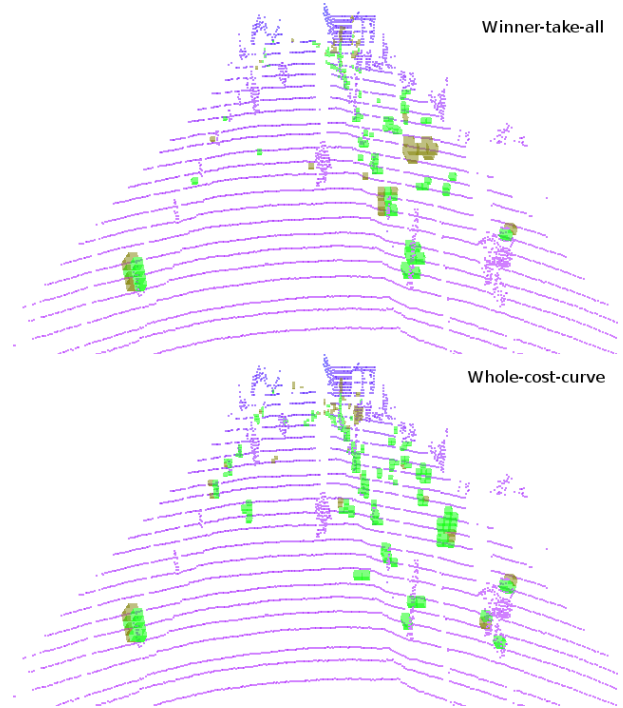

Fig. 8. Resulting occupancy grids when applying winner-take-all stereo (top) and our whole-cost-curve approach (bottom). Green areas are occupied cells that match ground-truth data, brown areas are mistakes. A close-up of the map is shown, for clarity.

The image shown for the winner-take-all approach on Fig. 8 was obtained using a Peak Ratio threshold equal to 3. This was the value we found lead to the best results. The result of applying a lower threshold, 1.5, is shown on Fig. 9.
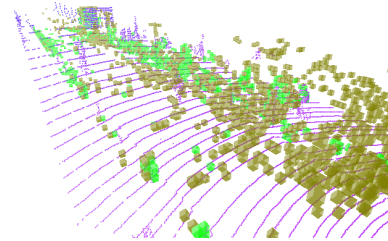

Fig. 9. Resulting grid for the winner-take-all approach when using low filtering thresholds (Peak Ratio 1.5). Brown cells are mistakes.

To quantitatively evaluate performance of the occupancy grid methods we take two measures: number of $hits$ or number of occupied cells matching ground-truth data; and the number of $mistakes$ or false-positives.

We run the occupancy grid methods on the two mentioned KITTI datasets. The winner-take-all method was run with

different values of the Peak Ratio threshold (from 1.5 to 12 in intervals of 0.5). Performance results are shown in Fig. 10. In both datasets, the whole-cost-curve approach achieves less $mistakes$ for the same number of $hits$ when compared to the winner-take-all approach. A high number of correct occupancy assignments ($hits$) is still achieved.
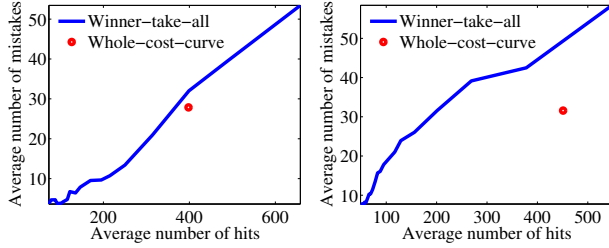

Fig. 10. Performance of the implemented occupancy grid with whole-cost-curve and winner-take-all stereo. Winner-take-all results were taken for different filtering thresholds (Peak Ratio). Left: Results from dataset "2011-09-28 drive 0038". Right: dataset "2011-09-28 drive 0045".

By applying less filtering to the stereo matches, a winner-take-all approach can actually capture more $hits$ than our approach. This comes, as already seen on the Virtual Environment experiment at the cost of a high number of $mistakes$. Both Fig. 9 and the blue curve in Fig. 10 also illustrate this property. However, if one can find post-filtering methods able to deal with high error-rates then a traditional approach could still be suitable.

C. Computation time

Finally we evaluate the proposed approach in terms of computation time. We compare computation times for different image sizes in Table I. Complexity of the sensor model computation is the same whether a whole-cost-curve or winner-take-all approach is used. Winner-take-all computes a maximum from the cost-curve, our approach uses confidence measure (7) on that same curve. The added processing time comes from the computation of exponentials ($exp(x)$) and divisions of the proposed confidence measure.

TABLE I
COMPUTATION TIMES (S)

Image Size	Task	Winner-take-all	Whole-cost-curve
320x240	Stereo Matching	0.04	0.04
	Inv. sensor model	0.02	0.06
640x480	Stereo Matching	0.20	0.20
	Inv. sensor model	0.12	0.16

Computation times were measured on a Intel i7, 2.7Ghz single core execution, without any floating point optimization or exponential function approximation. The proposed whole-cost-curve formulation of the occupancy grid predictably requires additional time for the inverse sensor model computation due to exponentiation of all costs. Nevertheless, the extra time is still not prohibitive for regular sized images.

V. CONCLUSIONS

An occupancy grid formulation integrating the whole cost-curve of stereo matching was proposed and evaluated experimentally. We concluded that the proposed approach generally leads to a high level of 3D reconstruction and low error ratios both in virtual and challenging outdoors environments.

We also tested the behaviour of a typical winner-take-all approach with different filtering levels only to find out that a trade-off between holes in the reconstruction and false-positives is necessary. That trade-off is easier to manage in the proposed approach, since better final maps are obtained (Fig. 5 and 8) without the requirement of parameter tuning.

From the simulation experiment we concluded that the proposed approach better deals with repetitive patterns that lead to uncertainty. Whole-cost-curve integration brings more evidence to the right matches, eventually leading to better reconstruction: without pre or post discarding of any matches. Depending on the stereo confidence measure used, the proposed approach is slightly more computationally expensive than a winner-take-all one, in the order of tens of milliseconds.

This paper shows the advantages of integration of all information returned by stereo into occupancy grids. For future work we will pursue new and more effective stereo confidence measures and strategies for improving occupancy grid results while maintaining low error rates. The free-space model will also be explored.

REFERENCES

[1] A. Elfes, "Using occupancy grids for mobile robot perception and navigation," *Computer*, vol. 22, no. 6, pp. 46–57, 1989.
[2] L. Matthies and A. Elfes, "Integration of sonar and stereo range data using a grid-based representation," in *IEEE International Conference on Robotics and Automation*. IEEE, 1988, pp. 727–733.
[3] D. Murray and J. J. Little, "Using real-time stereo vision for mobile robot navigation," *Autonomous Robots*, vol. 8, no. 2, pp. 161–171, 2000.
[4] F. Andert, "Drawing stereo disparity images into occupancy grids: measurement model and fast implementation," in *IEEE International Conference on Intelligent Robots and Systems*, 2009, pp. 5191–5197.
[5] M. Perrollaz, J.-D. Yoder, A. Spalanzani, C. Laugier, *et al.*, "Using the disparity space to compute occupancy grids from stereo-vision," in *International Conference on Intelligent Robots and Systems*, 2010.
[6] A. Suppes, F. Suhling, and M. Hötter, "Robust obstacle detection from stereoscopic image sequences using kalman filtering," *Pattern Recognition*, pp. 385–391, 2001.
[7] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 3354–3361.
[8] A. Elfes, "Sonar-based real-world mapping and navigation," *IEEE J. Robot. Automat.*, vol. 3, no. 3, pp. 249–265, 1987.
[9] K. Konolige, "Improved occupancy grids for map building," *Autonomous Robots*, vol. 4, no. 4, pp. 351–367, 1997.
[10] R. Shade and P. Newman, "Choosing where to go: Complete 3d exploration with stereo," in *IEEE International Conference on Robotics and Automation*, 2011, pp. 2806–2811.
[11] L. Matthies and P. Grandjean, "Stochastic performance, modeling and evaluation of obstacle detectability with imaging range sensors," *IEEE Trans. Robot. Automat.*, vol. 10, no. 6, pp. 783–792, 1994.
[12] X. Hu and P. Mordohai, "A quantitative evaluation of confidence measures for stereo vision," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 34, no. 11, pp. 2121–2133, 2012.
[13] L. Matthies, M. Okutomi, *et al.*, "A bayesian foundation for active stereo vision," *Proc. SPIE Sensor Fusion II: Human Machine Strategies*, pp. 62–74, 1989.
[14] D. Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, 2003, pp. I–195.