

Experiments with a New Area-Based Stereo Algorithm

A. Fusiello¹, V. Roberto¹, and E. Trucco²

¹ Machine Vision Laboratory, Dept. of Informatics
University of Udine, Italy

² Dept. of Computing and Electrical Engineering
Heriot-Watt University, UK

Abstract. We present a new, efficient stereo algorithm addressing robust disparity estimation in the presence of occlusions. The algorithm uses multiple windows and left-right consistency to compute disparity and its associated uncertainty. We demonstrate and discuss performances with both synthetic and real stereo pairs, and show how our results improve on those of closely related techniques for both robustness and efficiency.

1 Introduction

The aim of computational stereopsis is to reconstruct the 3-D geometry of a scene from two (or more) views, which we call *left* and *right*, taken by pinhole cameras (for a comprehensive review on computational stereo, see [8]). A well-known problem is *correspondence*, i.e., finding which points in the left and right images are projections of the same scene point (*a conjugate pair*). This is approached as search: finding the element in the right image which is most similar, according to a similarity metric, to a given element in the left image (a point, region, or generic feature).

Area-based (or *correlation-based*) algorithms [1,3] match small image windows centered at a given pixel, assuming that the gray levels are similar. They yield dense depth maps, but fail within occluded areas and poorly textured regions. *Feature-based* algorithms [6,12] match local cues (e.g., edges, lines, corners) and provide robust, but sparse, disparity maps requiring interpolation. These algorithms rely on feature extraction.

Several factors make the correspondence problem difficult: (i) its inherent *ambiguity* requires the introduction of physical and geometrical constraints; (ii) *occlusions*, i.e., points in one image with no corresponding point in the other; (iii) *photometric distortions* [2] arising when conjugate pair pixels have significantly different intensities; and (iv) *figural distortion* [9] that makes the projected shapes different in the two images.

This paper presents a new robust area-based algorithm, addressing all problems (i)-(iv) listed above by exploiting symmetry in matching and multiple windows. For this reason it will be called *Symmetric Multi-Window* (SMW) algo-

rithm in the following. Preliminaries needed to meet some assumptions on image pairs is first illustrated (Sect. 2). The SSD correlation method is presented (Sect. 3), followed by our adaptive, multi-window scheme (Sect. 4), which contrasts distortions and yields accurate disparities. Robust disparity estimates in the presence of occlusions are achieved thanks to the *left-right consistency constraint* (Sect. 5); the associate uncertainty is estimated too (Sect. 6). SMW algorithm implementation is sketched in Sect. 7. Experimental results are presented in Sect.s 8 and 9.

2 Assumptions

Our algorithm for disparity computation assumes that conjugate pairs lie along raster lines. In general this is not true, therefore stereo pairs need to be *rectified* – after appropriate camera calibration – to achieve epipolar lines parallel and horizontal in each image [5].

The SMW algorithm also assumes that the image intensity of a 3D point is the same on the two images. If this is not true, the images must be *normalised*. This is done by a simple algorithm [2] which computes the parameters of the gray-level transformation

$$I_l(x, y) = \alpha I_r(x, y) + \beta \quad \forall (x, y)$$

by fitting a straight line to the plot of the left cumulative histogram versus the right cumulative histogram.

3 Solving Correspondence

Similarity scores are computed, for each pixel in the left image, by comparing a fixed small window centered on the pixel with a window in the right image, shifting along the raster line. As a similarity measure we adopt the Euclidean distance, which is also called SSD (*Sum of Squared Differences*) error:

$$C(x, y, d) = \frac{\sum_{(\xi, \eta)} [I_l(x + \xi, y + \eta) - I_r(x + \xi + d, y + \eta)]^2}{\sqrt{\sum_{(\xi, \eta)} I_l(x + \xi, y + \eta)^2 \sum_{(\xi, \eta)} I_r(x + \xi + d, y + \eta)^2}} \quad (1)$$

where $\xi \in [-n, n]$, $\eta \in [-m, m]$. The computed disparity is the one that minimises the SSD error. *Subpixel precision* is achieved by fitting a curve to the errors in the neighbourhood of the minimum [1].

If one computes SSD by a straightforward implementation of (1), the asymptotic complexity of the resulting algorithm is $O(N^2nm)$, with N the image size. However, one should observe that squared differences need to be computed only once for each disparity and that the sum over the window should not be recomputed from scratch when it is moved by one pixel. The optimised implementation that follows from this observation [3] has a computational complexity of $O(4N^2)$, which is independent of the window size.

4 Window Shaping

As observed by Kanade and Okutomi [9], when the correlation window covers a region with non-constant disparity, area-based matching is likely to fail, and the error in the depth estimates grows with the window size. Reducing the latter, on the other hand, makes the computed disparity more noise-sensitive. To overcome such difficulties, Kanade and Okutomi proposed a statistically sound, adaptive technique which selects at each pixel the window size that minimises the uncertainty in the disparity estimates.

In this work we take the multiple window approach in the simplified version proposed by [7]. For each pixel we perform the correlation with nine 7×7 different windows (shown in Fig. 1), and retain the disparity with the smallest SSD error value. The idea is that a window yielding a smaller SSD error is more likely to cover a constant depth region; in this way, *the disparity profile itself drives the selection of an appropriate window*.

5 Left-Right Consistency

Occlusions create points that do not belong to any conjugate pairs. In many cases, occlusions occur at depth discontinuities: indeed, one may observe that occlusions on one image correspond to disparity jumps on the other. Although occlusions help the human visual system in detecting object boundaries, in computational stereo they are a major source of errors.

A key observation to address the occlusion problem is that *matching is not a symmetric process*: when searching for conjugate pairs, only the visible points in one image are matched. If the role of left and right images is reversed, new conjugate pairs are found. The so-called *left-right consistency constraint* [4] states that feasible conjugate pairs are those found with both direct and reverse matchings. Consider for instance an occluded point, e.g., B in the left image of Fig. 2: although it has no corresponding point in the right image, the SSD minimisation matches it to some point (C') anyhow. One can see that the latter point, in turn, corresponds to a different point in the left image, but this information is available only by searching from right to left.

In our approach, occlusions are detected by checking the left-right consistency, and suppressing unfeasible matches accordingly. For each point (x, y) on the left image the disparity $d_l(x, y)$ is computed as described in Sect. 2. The process is repeated after reversing the two images, in order to compute $d_r(., .)$. If $d_l(x, y) = -d_r(x + d_l(x, y), y)$ the point keeps its computed left disparity, otherwise it is marked as occluded and a disparity is assigned heuristically: following [10], we assume that occluded areas, occurring between two planes at different depth, take the disparity of the deeper plane.

6 Uncertainty Estimates

Area-based algorithms are likely to fail not only in occluded regions, but also in poorly textured regions, which make disparity estimates more uncertain. It is

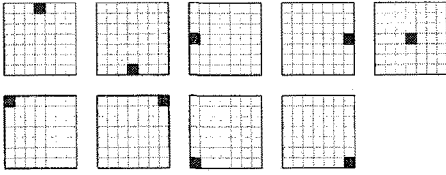


Fig. 1. The nine asymmetric correlation windows. The pixel for which disparity is computed is highlighted.

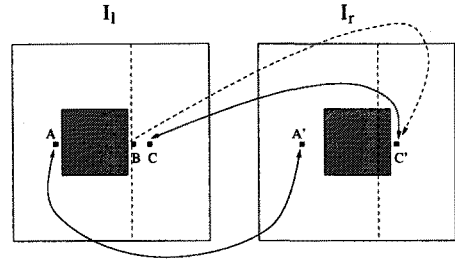


Fig. 2. Left-right consistency. Point B is given C' as a match, but C' matches $C \neq B$. The pair (B, C') can be suppressed.

therefore essential to compute confidence measures for disparities, which enables one to fill in gaps of the depth maps by fusing multiple views. Several techniques are available to estimate uncertainty, most of them based on the shape of the SSD error function [1,11,13].

In our approach we take advantage of the fact that disparity values computed with different windows are sensitive to the signal-to-noise ratio (SNR): as the latter decreases, the variance of the disparity values increases (see Fig. 8). Hence, we take it as an uncertainty measure for the computed disparity; occluded points are assigned infinite uncertainty.

7 The SMW Algorithm

The SMW algorithm can be implemented in the following steps (disparity is assumed to be positive, that is the right view is right-shifted with respect to the left view):

1. Compute disparity values with SSD correlation from left to right, using the asymmetric windows, and retain the lowest SSD disparity.
2. Compute uncertainty as the variance of disparity values.
3. Do Step 1 by reversing left and right images.
4. Check the left-right consistency and suppress matches accordingly. Unmatched pixels are marked as occluded.
5. Compute subpixel refinement of disparity values.
6. Set to infinite the uncertainty of occluded pixels and fill occluded regions with disparity values from right to left.

To facilitate reimplementations and experiments with the SMW, the C code of the algorithm is available via anonymous ftp at taras.dimi.uniud.it/pub/code/smw.tar.gz

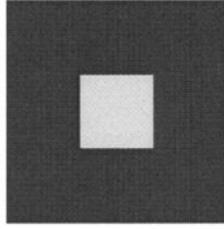
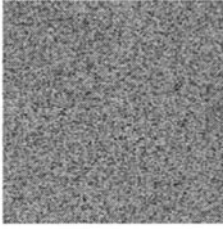


Fig. 3. Square random-dot stereogram. The left image of the stereogram is shown (left). The right one is computed by warping the latter with a given disparity pattern (right). The square has disparity 10 pixel, the background 3 pixel.

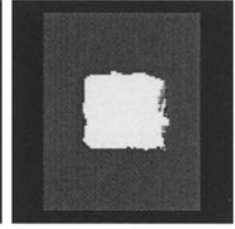
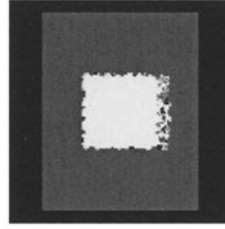


Fig. 4. Computed disparity map by SSD correlation for the square random-dot stereogram in Fig. 5 with 3×3 window (left) and 7×7 window (right); MAE is 0.240 and 0.144, respectively.

8 Experiments with Synthetic Data

We first performed experiments on uncorrupted random-dot stereograms (Fig. 3), in order to assess the algorithm in a simple, albeit not trivial, case. Disparity maps are gray-level encoded (the brighter the closer). Images have been equalised to improve readability, subpixel-accuracy values have been computed and rounded to integers. The estimated Mean Absolute Error (MAE), that is the mean of the absolute value of differences between computed disparity and ground true disparity, was computed.

Simple SSD correlation applied to random-dot stereograms shows how most of the problems outlined in the previous sections affect the disparity computation. Fig. 4 shows the disparity maps computed with the SSD correlation algorithm, with fixed 3×3 and 7×7 windows. In both pictures it is visible the effect of disparity jumps (near the left and horizontal borders of the square patch) and occlusions (near the right border of the square patch).

The SMW algorithm was applied to the square random-dot stereograms of Fig. 3 and to a circular random-dot stereogram, not shown here. Fig. 5 and Fig. 6 show the disparity maps computed by SMW and the estimated uncertainty maps (the darker the lower) in both cases. The estimated MAE is negligible and may be ascribed to the subpixel accuracy only. The occluded points, shown in white in the uncertainty maps are recovered with 100% accuracy, in both cases. The circle random-dot stereogram shows that the algorithm is not biased toward square disparity patterns, as it may seem due to the shape of the windows. The reader may compare the present results to those reported in [2].

As for efficiency, running on a SUN SparcStation 4 - 110MHz under SunOS 5.5, the SMW algorithm takes 8 seconds, on the average, to compute the depth maps on 128×128 input images. Although accuracy results are comparable to those of closely related techniques, such as [9], the efficiency of SMW is clearly superior.

Further experiments with noisy random-dot stereograms show a graceful degradation when noise increases. Gaussian noise with zero mean and increasing variance was added independently to both images of the square random-dot stereogram. Fig. 7 shows the MAE vs noise standard deviation for SMW and SSD correlation. Each point depicts the average result of 20 independent trials.

In order to assess the uncertainty map produced by SMW, the average uncertainty computed over a square patch of uniform disparity was plotted against the SNR (Fig. 8). The plot shows that the computed uncertainty consistently increases as the SNR decreases.

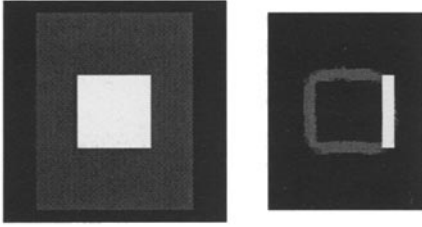


Fig. 5. Computed disparity map (left) by SMW for the square random-dot stereogram and its uncertainty (right). MAE is 0.019.

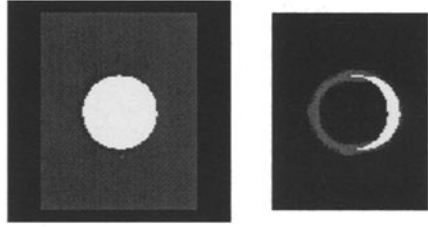


Fig. 6. Computed disparity map (left) by SMW for the circle random-dot stereogram and its uncertainty (right). MAE is 0.026.

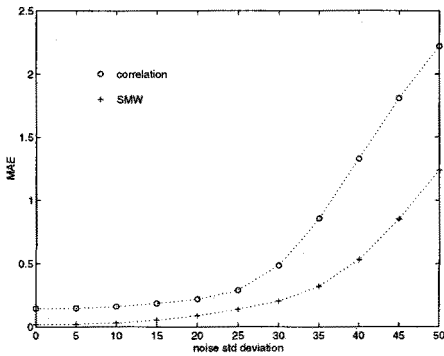


Fig. 7. MAE vs noise standard deviation for the square random-dot stereogram. Window size is 7×7 .

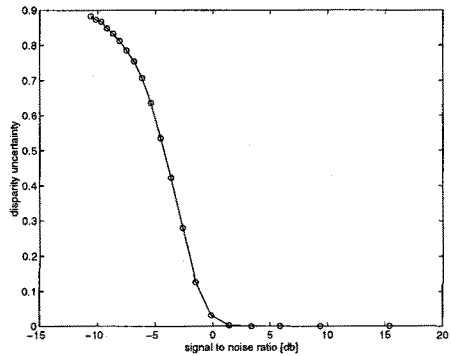


Fig. 8. Mean uncertainty vs SNR for a constant disparity region of the square random-dot stereogram.

9 Experiments with Real Data

We performed experiments on standard image pairs from the JISCT (JPL-INRIA-SRI-CMU-TELEOS) stereo test set. Only the “Parking meter”(Fig 9) is reported here for reason of space.

Small values cannot be appreciated in spite of histogram equalisation, due to the large difference between high-uncertainty occlusion points and the rest of the image. Although a quantitative comparison with published results was not possible with real images, the quality of SMW results seems perfectly comparable to that of results reported, for example, in [14,2].

Running on the same hw/sw platform, our current implementation takes 50 seconds, on the average, to compute depth maps from 256×256 pairs, and a disparity range of 10 pixels.

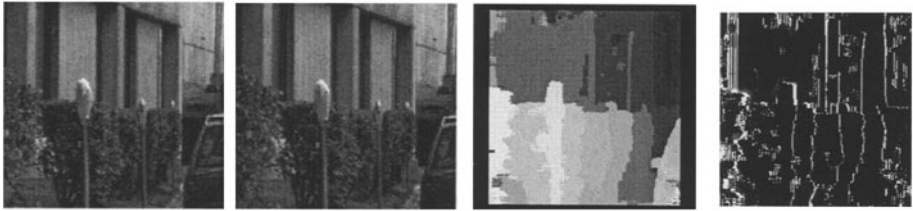


Fig. 9. The “Parking meter” stereo pair; the disparity (left) and uncertainty maps (right).

10 Conclusions

We have introduced SMW, a new, efficient algorithm for stereo reconstruction, based on a multi-window approach, and taking advantage of left-right consistency. Our tests have shown the advantages offered by SMW. The adaptive, multi-window scheme yields robust disparity estimates in the presence of occlusions, and clearly outperforms fixed-window schemes.

The left-right consistency check proves very effective in eliminating false matches and identifying occluded regions (notice that this can be regarded as a segmentation method in itself). In addition, disparity is assigned to occluded points heuristically, thereby achieving reasonable depth maps even in occluded areas. Uncertainty maps are also computed, allowing the use of SMW as a module within more complex data fusion frameworks [13]. Areas of lower SNR are consistently marked with higher uncertainty.

The main disadvantage is that the window size remains a free parameter; we are considering a multi-resolution extension to the SMW algorithm, where correlation is performed with a 3×3 window at different resolution levels.

Work is in progress also to embed the SMW module in a dynamic stereo fusion system.

Acknowledgements

This work was partially supported by a British Council-MURST/CRUI grant.

References

1. P. Anandan. A computational framework and an algorithm for the measurement of visual motion. *International Journal of Computer Vision*, 2:283–310, 1989.
2. I. J. Cox, S. Hingorani, B. M. Maggs, and S. B. Rao. A maximum likelihood stereo algorithm. *Computer Vision and Image Understanding*, 63(3):542–567, May 1996.
3. O. Faugeras, B. Hotz, H. Mathieu, T. Viéville, Z. Zhang, P. Fua, E. Théron, L. Moll, G. Berry, J. Vuillemin, P. Bertin, and C. Proy. Real-time correlation-based stereo: algorithm, implementation and applications. Technical Report 2013, Unité de recherche INRIA Sophia-Antipolis, Août 1993.
4. P. Fua. Combining stereo and monocular information to compute dense depth maps that preserve depth discontinuities. In *Proceedings of the International Joint Conference on Artificial Intelligence*, Sydney, Australia, August 1991.
5. A. Fusiello, E. Trucco, and A. Verri. Rectification with unconstrained stereo geometry. Submitted to ICCV'98.
6. W.E.L. Grimson. Computational experiments with a feature based stereo algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(1):17–34, 1985.
7. S. S. Intille and A. F. Bobick. Disparity-space images and large occlusion stereo. In Jan-Olof Eklundh, editor, *European Conference on Computer Vision*, pages 179–186, Stockholm, Sweden, May 1994. Springer-Verlag.
8. M. R. M. Jenkin, A. D. Jepson, and J. K. Tsotsos. Techniques for disparity measurements. *CVGIP: Image Understanding*, 53(1):14–30, 1991.
9. T. Kanade and M. Okutomi. A stereo matching algorithm with an adaptive window: Theory and experiments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(9):920–932, September 1994.
10. J. J. Little and W. E. Gillett. Direct evidence for occlusions in stereo and motion. *Image and Vision Computing*, 8(4):328–340, 1990.
11. L. Matthies, T. Kanade, and R. Szelisky. Kalman filter based algorithms for estimating depth from image sequences. *International Journal of Computer Vision*, 3:209–236, 1989.
12. Y. Ohta and T. Kanade. Stereo by intra- and inter-scanline search using dynamic programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(2):139–154, 1985.
13. E. Trucco, V. Roberto, S. Tinonin, and M. Corbatto. SSD disparity estimation for dynamic stereo. In R. B. Fisher and E. Trucco, editors, *Proceedings of the British Machine Vision Conference*, pages 342–352. BMVA Press, 1996.
14. Y. Yang and A. L. Yuille. Multilevel enhancement and detection of stereo disparity surfaces. *Artificial Intelligence*, 78(1-2):121–145, October 1995.