Parallel Multiscale Stereo Matching Using Adaptive Smoothing [1]

Jer-Sen Chen [2] *and Gérard Medioni*
Institute for Robotics and Intelligent Systems
University of Southern California
Los Angeles, California 90089-0273
Email: medioni@usc.edu

We present a multiscale stereo algorithm whose design makes it easily implemented on a fine grain parallel architecture. The matching at a given scale is performed in accordance with the algorithm proposed by Drumheller and Poggio, using edges as primitives. Using multiple scales is in general difficult and costly because of the correspondence problem between scales, but this problem is solved here using adaptive smoothing, a process which preserves edge location at different scales. The results are better than those obtained at a single scale, and the algorithm runs in a few seconds on a Connection Machine.

Introduction

Range or depth information has long been considered essential in the analysis of shape. High-resolution range information can be obtained directly so long as the range sensor is available. Binocular stereo has been widely used to extract the range information when such a high-resolution sensor is absent. Barnard and Fischler [Barnard82] define six steps necessary to stereo analysis: image acquisition, camera modeling, feature acquisition, image matching, depth reconstruction and interpolation. Multiple scale processing not only provides a description of the signal but only facilitate a coarse-to-fine hierarchical processing for various vision tasks. The correspondence problem in stereo matching is usually very tedious and can be alleviated through a multiple scale approach.

The correspondence problem between two images can be solved by matching specific features such as edges, or by matching small regions by the correlation of the image intensities. Edgel-based stereo matching techniques usually use the edges characterized by the derivatives of a smoothed version of the signal, for instance the zero-crossings of a Laplacian-of-Gaussian convolved image. The correlation-based stereo matching measures the correlation of the image intensity patches centered around the matched pixels. Our multiscale stereo matching is edgel-based and the matching primitives are edgels extracted with adaptive smoothing. Since adaptive smoothing provides accurate edge detection across different scales, it facilitates a straightforward multiscale stereo matching.

To identify corresponding locations between two stereo images, or among a sequence of motion images, is difficult because of the *false targets problem.* Certain constraints and assumptions have to be made in order to establish the correct pairings. The *Uniqueness constraint* [Marr82] states that there is at most one match from each line-of-sight since the depth value associated with each matching primitive, left or right, must be unique. The *Continuity constraint* [Marr82] states that the depth map of an image should be mostly continuous, except at those points where depth discontinuities occur. Therefore neighboring potential matches having similar disparity values should support each other. The *Opacity constraint* further limits the occurrence of the false targets. Extending the uniqueness constraint, which limits to only one match along each line-of-sight, the opacity constraint states that there is at most one match in the hourglass-shape forbidden zone bounded by two lines-of-sight. The *Compatibility constraint* [Marr82] limits the construction of potential matches from matching primitives, for example, the potential matches are allowed to occur only when two zero-crossings from the **LoG** convolved images have the same sign. Further restrictions can be made on the orientation and the gradient of the matching edgels.

The false targets problem can be alleviated either by reducing the range and resolution of the disparity or by reducing the density of the matching features in the image. One commonly used

[1] This research was supported by the Defense Advanced Research Projects Agency under contract and F33615-87-C-1436, and monitored by Wright-Patterson Air Force Base.

[2] current address: Wright State University, Department of Computer Science and Engineering, Dayton, Ohio 45435 (email: jschen@odin.wright.edu)

method to obtain both resolution and range of disparity information is to apply a multi-resolution algorithm. The information obtained from the matching at coarse resolution can be used to guide the matching at fine resolution.

Marr and Grimson [Marr82,Grimson81] use Gaussian kernels as the scaling filters for multiscale stereo matching. The matches obtained at coarser scale establish a rough correspondence for the finer scales, thereby reducing the number of false matches, but a vergence control is necessary because of the poor accuracy of the **LoG** edge detection.

This is overcome here by using the adaptive smoothing, which preserves edge locations across scales. The next section briefly reviews the major properties of adaptive smoothing, and section 3 presents the stereo algorithm and results on two sets of images.

Adaptive Smoothing

We have recently introduced a formalism called adaptive smoothing [Saintmarc89]. It keeps all the desirable properties of Gaussian smoothing [Perona87,Chen89] and preserves the location of edges across scales, making the correspondence problem trivial. Adaptive smoothing, in its basic formulation, assumes that the signal is piecewise constant inside a region and uses an ideal step edge model. It not only preserves edges but also enhances them. With a suitable choice of the scale parameter, an accurate edge detection scheme at different scales can be achieved by adaptive smoothing and therefore facilitate multiple scale signal processing. A multiscale representation of the signal can be easily derived by choosing the necessary number of scales without dealing with the tedious correspondence problem as encountered in the traditional Gaussian scale space [Witkin83, Asada86]. Edges corresponding to adaptive smoothing are presented in the results section.

On the Connection Machine with 16K physical processors, adaptive smoothing takes about 11 milliseconds per iteration in the case of one pixel per physical processor, compared to 10 seconds per iteration on a serial machine (Symbolics 3645) for a 128×128 image.

Multiscale Stereo

As mentioned in the introduction, multiscale processing is often used in a coarse to fine strategy to solve the matching problem, but one must solve the correspondence problem between scales. Adaptive smoothing can overcome this disadvantage with its accuracy of edges over scales. The matching results at coarser scale with adaptive smoothing are therefore much more reliable and the propagation of the disparity information between scales is straightforward.

We have used our adaptive smoothing and implemented a multiscale stereo matching algorithm to extract the matching features. It is based on Drumheller and Poggio's [Drumheller86] parallel stereo matching implementation on the Connection Machine [Hillis85]. The parallel stereo matching, as in most stereo matching algorithms, utilizes the *uniqueness* the *continuity* constraint on the surface and therefore the values on the disparity map. It also imposes the *opacity* constraint on the surfaces and the *compatibility* constraint on the matching of the edges.

We use three scales, namely coarse, intermediate and fine, in our multiscale stereo matching. We first extract edges at coarse scale for both images using adaptive smoothing. The stereo images are assumed to be epipolarly registered and the matching is performed scan-line by scan-line. A potential match is marked only when the corresponding edges from the two images have approximately the same orientation and gradient. Imposing the continuity constraint, the number of potential matches is counted over a flat uniformly-weighted square support (chosen for computational convenience) centered at each pixel. Enforcing the opacity and uniqueness constraints, there must be no more than one match in the forbidden zone, therefore a *winner-take-all* strategy is applied in the forbidden zone.

The matches at the edge locations from the coarse scale are then propagated to the intermediate scale. Each match at coarse scale generates a forbidden zone which forbids potential matches to be marked at intermediate scale. This greatly reduces the number of potential matches at intermediate scale and therefore facilitates producing more reliable matching results. After potential matches are formed at intermediate scale, the same continuity, opacity and uniqueness constraints are employed

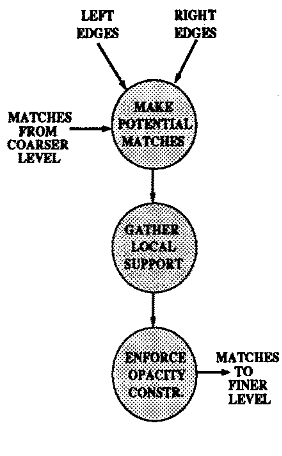

Figure 1: Flowchart of Multiscale Stereo Matching

to produce the matches which are then propagated to the fine scale. A flow chart of the process is shown in figure 1.

The multiscale stereo matching algorithm is implemented on the Connection Machine. The accurate edge detection by adaptive smoothing provides a very simple control mechanism for multiscale processing, and the simplicity is essential when considering parallel implementation.

We first show in figure 2 stereo matching of the aerial view of Pentagon ($256 \times 256 \times 8bits$, courtesy of Dr. W. Hoff). The pair is epipolarly registered, i.e. the two images correspond to each other scan-line by scan-line. Subfigures (a) and (b) show the left and right view of the original stereo pair. Part (c) shows the matching result using single (fine) scale. Subfigures (d), (e) and (f) show the left edgels, right edgels and the matching result at the coarse scale, while (g), (h) and (i) are for intermediate scale and (j), (k) and (l) are for fine scale. The scaling parameter k is set to 8, 4 and 0 (no smoothing) for the coarse, intermediate and fine scales respectively. The range of disparity for this example is -5 (farther) to 5 (nearer) pixels, and the brighter gray level values indicate a position closer to the viewer.

We show in figure 3 another example of a fruit scene ($256 \times 256 \times 8bits$, courtesy of Dr. T. Kanade, Carnegie-Mellon University). The scaling parameter k is set to 12, 6 and 0 for coarse, intermediate and fine scales respectively. The range of disparity is -15 to 22 pixels. Figures 3(a) and (b) show the stereo image pair. The results of single scale and multiscale matching are shown in figure 3(c) and (d) respectively.

Table 1 summarizes the statistics of the number of potential matches for the stereo matching. The column "multiple" stands for multiscale stereo matching and column "single" stands for stereo matching at each scale individually. Since the matches at coarser scale are used to form the forbidden zones when constructing potential matches at finer scale, the number of potential matches is greatly reduced at the finer scales as we can observe from the table.

Conclusion

We have shown a multiscale coarse-to-fine hierarchical matching of stereo pairs which uses adaptive smoothing to extract the matching primitives. The number of matching primitives at coarse scale is small, therefore reducing the number of potential matches, which in return increases the reliability of the matching results. A dense disparity can be obtained at a fine scale where the density of edgels is very high. The control strategy is very simple compared to other multiscale approaches such as the ones using Gaussian scale space, this results from the accuracy of edges detected by

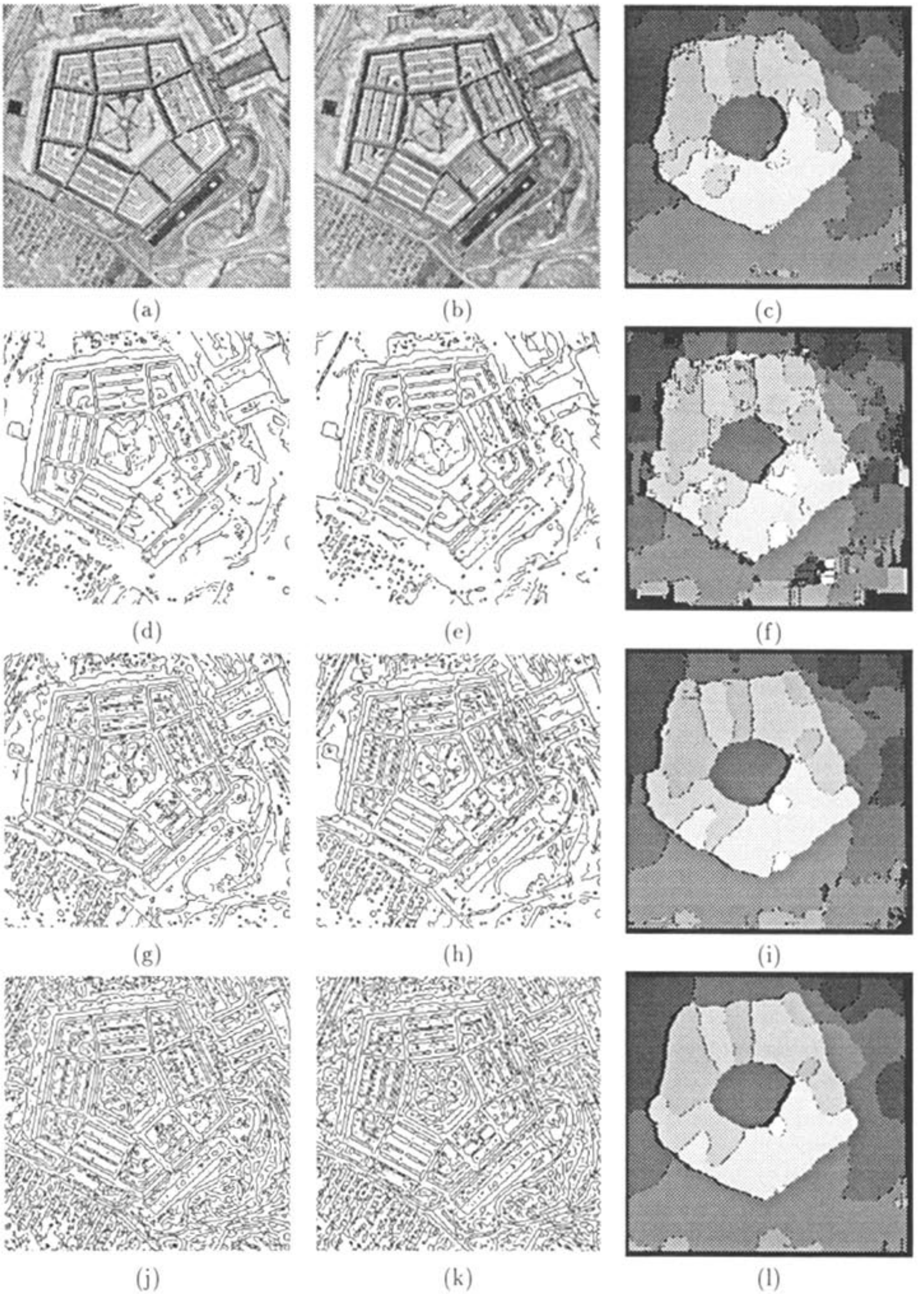

Figure 2: Stereo Matching of the Pentagon

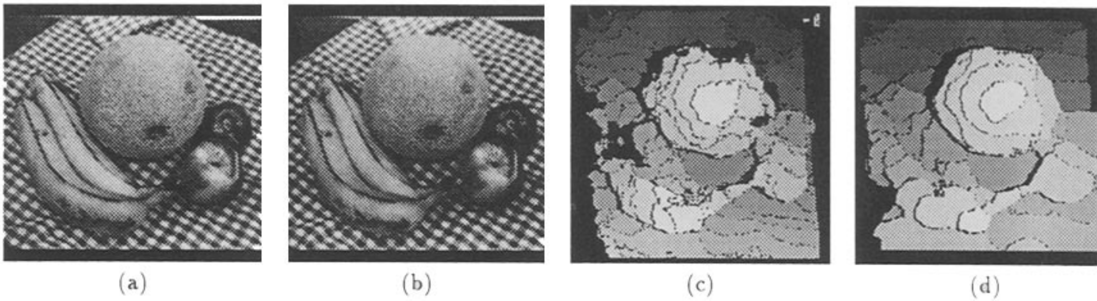

Figure 3: Stereo Matching of the Fruit Scene

	Pentagon			Fruit		
	coarse	Interm.	fine	coarse	Interm.	fine
singlescale	22670	28892	33323	36885	80045	120026
multiscale	22670	11372	11279	36885	37588	35676

Table 1: Number of Potential Matches

adaptive smoothing at different scales. The simplicity of the control strategy is especially important for low-level processing, and makes parallel implementation quite simple.

References

[Asada86] H. Asada and M. Brady. "The Curvature Primal Sketch", *IEEE Tran. on Pattern Analysis and Machine Intelligence*, Vol. 8, No. 1, January 1986, pp 2-14.

[Barnard82] Barnard, S. and Fischler, M., "Computational Stereo", *ACM Computing Surveys*, Vol. 14, No. 4, December 1982, pp 553-572.

[Burt83] Burt, P.J.,"Fast Algorithm for Estimating Local Image Properties," *Journal of Computer Vision, Graphics, Image Processing*, Vol. 21, March 1983, pp. 368-382.

[Chen89] Chen, J.S., "Accurate Edge Detection for Multiple Scale Processing", Ph.D. thesis, University of Southern California, October 1989.

[Drumheller86] Drumheller, M., Poggio, T., "On Parallel Stereo", *Proceedings of the 1986 IEEE International Conference on Robotics and Automation*, April 1986, pp 1439-1488.

[Grimson81] Grimson, W.E.L., *From Images to Surfaces*, MIT Press, Cambridge, USA, 1981.

[Hillis85] D. Hillis. *The Connection Machine.* MIT Press, Cambridge, MA, 1985.

[Marr82] Marr, D., *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* W. H. Freeman and Co., San Francisco, 1982.

[Perona87] Perona, P., Malik, J., "Scale Space and Edge Detection using Anisotropic Diffusion", *Proceedings of the IEEE Workshop on Computer Vision*, Miami Beach, Fl. 1987, pp 16-22.

[Saintmarc89] Saint-Marc, P., Chen J.S. and Medioni, G., "Adaptive Smoothing: A General Tool for Early Vision", *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, San Diego, June 1989.

[Witkin83] Witkin, A.P., "Scale Space Filtering", *Proceedings of International Joint Conference on Artificial Intelligence*, pp 1019-1022, Karlsruhe, 1983.