TR－A－0129

Enhanced Discontinuity Detection from Postulated Discontinuities

E. B. Gamble Jr.

1992. 1.13
（1992. 1.10受付）

ATR視聴覚機構研究所

〒619-02 京都府相楽郡精華町光台2-2　☎07749-5-1411

ATR Auditory and Visual Perception Research Laboratories

2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-7749-5-1411
Facsimile: +81-7749-5-1408
Telex: 5452-516 ATR J

Contents

Abstract

We describe a discontinuity detector that integrates visual information to suppress the pervasive noise in surface property data. The discontinuity detector is based on the notion of 'constrained kernels' and 'postulated discontinuities.' The kernels arise as solutions to the diffusion equation in the presence of *local, static* boundary conditions provided by the postulated discontinuities. The resulting kernels smooth the surface property data; they are Gaussian-like except near the postulated discontinuities. Unlike most visual integration schemes, our detector does not suffer from problems regarding computability and parameter specification; it is fast and accurate. In addition, the detector eliminates the pervasive 'displacement errors' in surface property data. We describe these displacement errors, integrate intensity edges (as postulated discontinuities) with stereo depth and optical-flow to compute depth and motion discontinuities, and compare our detector to other discontinuity detection approaches.

This is a reformated version of a paper submitted to the Second European Conference on Computer Vision in Santa Margherita Ligure, Italy. The original submission to ECCV'92 was dated 8 October 1991.

1 Introduction

The problem of discontinuity detection is of fundamental importance in computer vision. A discontinuity represents a location in an image where some property of the surface in the imaged scene, such as depth or motion, changes abruptly. Such changes in surface property are important because, under some general assumptions, they correspond to boundaries between objects. Unfortunately, the detection of discontinuities is a difficult task because the surface properties computed by early vision algorithms are fraught with errors near the discontinuities themselves[1, 2].

Somewhat surprisingly, few discontinuity detectors adopt a noise model consistent with the errors in surface property data near the discontinuities. Typically a spatially white, Gaussian noise process is assumed. However, as we illustrate in the next section, the noise is not spatially white; the surface property data is plagued by *displacement errors* near the actual, physical discontinuities. Based solely on surface property data there is no *a priori* method to identify these displacement errors – additional information is required. Consequently, if a discontinuity detector does not model these displacement errors the detector will mark discontinuities as *displaced* from the physical discontinuity; clearly this is an error.

Recently, studies on integration of visual information have shown that discontinuity detection can be improved if prior, though uncertain, evidence concerning the discontinuity *location* exists[3]. This prior information properly locates the discontinuity in spite of the displacement errors and then only the presence or absence of the discontinuity need be ascertained. We extend this recent work by providing a major reformulation and simplification of the discontinuity detector and also by definitively describing the displacement errors and their characteristics.

The solution to the diffusion equation in the presence of *postulated discontinuities* provides the basis for our discontinuity detector formulation. The postulated discontinuities provide the additional, local information needed to eliminate displacement errors and they also limit the surface reconstruction problem significantly enough to allow for efficient implementation. Essentially, the postulated discontinuities act as mirror boundary conditions for the diffusion equation solution and thereby inhibit the smoothing between neighboring lattice sites when one of the sites coincides with a postulated discontinuity. These local boundary conditions lead to Gaussian-like but slightly deformed *constrained kernels*.

Most discontinuity detectors, whether based on integration of visual information or not, simultaneously seek to smooth the surface properties and to detect the discontinuities (refer to Section 4). We will argue that given the nature of displacement errors it is inappropriate to simultaneously smooth the surface while detecting discontinuities; the characteristics of displacement errors require their removal prior to discontinuity detection. Consequently, discontinuity detection can be considered as a variant to traditional edge detection[4, 5]; that is, as a sequential process of

convolution to eliminate the dominant noise followed by hysteresis on the gradient magnitude while suppressing neighbors to locate the discontinuity. The postulated discontinuities and resulting constrained kernels provide the means for removal of the displacement errors and permit formulation of the detector as a sequential process rather than as a simultaneous process.

Besides eliminating displacement errors, postulated discontinuities provide a convenient mechanism to integrate visual information. Multiple visual cues can be combined at the locations of postulated discontinuities thereby ensuring registration among the visual cues. Our detector integrates visual information by using intensity edges to postulate the locations of discontinuities. We will not discuss the issues of integration nor the use of intensity edges in visual integration in any detail; numerous sources exist[6, 7, 8, 3]. It is however important to note that the postulated discontinuities obviously cannot be the discontinuities. Ideally the postulated discontinuities are a *superset* of the discontinuities; although, the postulates are themselves noisy.

The organization of this article is as follows. Section 2 contains a discussion of the displacement errors in surface property data produced by early vision algorithms. This section includes a simple, working model for these errors. In Section 3 we construct the constrained kernels based on the postulated discontinuities and then illustrate the effectiveness of the kernels when applied to the displacement error model. In Section 4 the differences between our formulation and previous work on discontinuity detection are described. Finally, Section 5 presents discontinuity detection results from surface property data computed from real images.

2 Displacement Errors in Surface Property Data

The first step in our formulation is to describe the pervasive displacement errors. We show that, not unexpectedly, the noise is particularly bad near the discontinuities which we seek to detect. This observation mandates a discontinuity detector that uses prior information to postulate discontinuity locations.

The smoothness constraint that is needed to regularize the ill-posedness of many early vision computations[10] leads to displacement errors in surface property data. In order to examine these errors, we consider a typical early vision algorithm for the

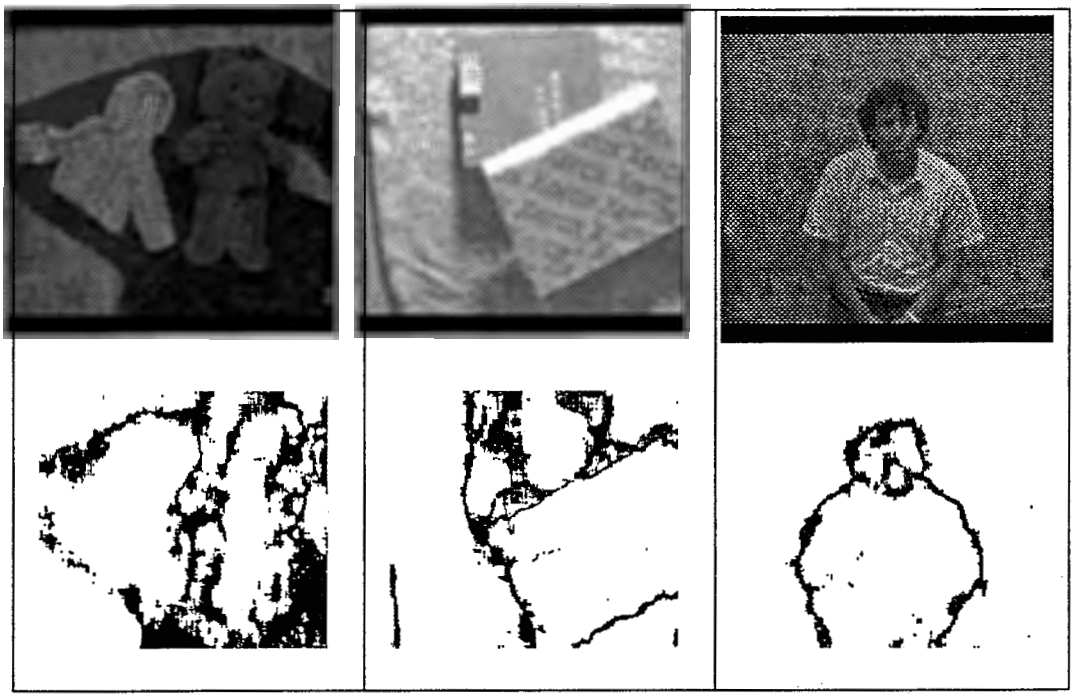

Figure 1: Examples of multimodal regions in correlation detectors[9, 2] are illustrated below the left image from a stereo pair. The multimodal regions are consistently near depth discontinuities.

computation of optical flow[11]. Assume that two images of the same scene exist, $E_1(\vec{r})$ and $E_2(\vec{r})$, where $\vec{r} = x\hat{x} + y\hat{y}$ is a vector in the image plane (with axes $\hat{x}$ and $\hat{y}$) and $E(\vec{r})$ is the intensity measured by the imaging device at pixel (x, y). The fundamental problem is to find the correspondence between each of the pixels in the two images. The x-y translation in image coordinates between the two corresponding pixels is the *disparity*, $\vec{d}(\vec{r})$, where $\vec{d}(\vec{r}) = d_x(\vec{r})\hat{x} + d_y(\vec{r})\hat{y}$. A solution for $\vec{d}$ is that function which yields an extremum in the following functional:

$$S() = \int \left\{ \phi \left[E_1(\vec{r}), E_2(\vec{r} + \vec{d}(\vec{r})) \right] + \lambda \|(\nabla \vec{d})\|^2 \right\} d\vec{r}. \tag{1}$$

Because of the smoothness constraint, $\vec{d}(\vec{r})$ must be a slowly varying function of $\vec{r}$. If $\vec{d}(\vec{r})$ is assumed constant over regions of size A, then Equation 1 simplifies to

$$S() = \int_A \phi \left[E_1(\vec{r}), E_2(\vec{r} + \vec{d}(\vec{r})) \right] d\vec{r} \tag{2}$$

where $S()$ is called the *match score*. The computation is a correlation in $\vec{d}(\vec{r})$ across $\vec{r}$; the solution is that $\vec{d}(\vec{r})$ which yields an extremum in Equation 2. The analysis for stereo is similar except that for the computation of $S()$, $\vec{r}$ is restricted to lie on an epipolar line.

The functional form of $S()$ yields important information about the matching process[9, 2]. If $S()$ is multimodal then, generally, the smoothness constraint has been violated. This occurs near discontinuities where part of region A correlates well at, say, d_1 but another part of A correlates well at d_2. Examples of multimodal correlation functions have been presented for optical flow[9] and stereo[2]. Figure 1 contains three images; each image is accompanied by a binary map that highlights those pixels where the correlation function is multimodal. These binary maps were computed with a pixel-based stereo algorithm[2] in which $\phi()$ of Equation 1 was the square of the difference between the intensity of each pixel in the stereo pair and $\vec{d}(\vec{r})$ was found as the minimum of Equation 2. The figure shows that the multimodal regions always occur near discontinuities in depth and, in some cases, in places unrelated to discontinuities.

Figures 2 and 3 illustrate displacement errors in stereo disparity. *Displacement* refers to one pixel near a discontinuity reporting the disparity of the pixels on the other side of the discontinuity; this displaced pixel is on the wrong side of the "actual" discontinuity. In Figure 2 the surface with disparity 2 and 3 is corrupted by 5 pixels in the middle-left of the figure with a disparity of 12. Figure 3 illustrates stereo disparity computed for several different patch sizes A (from Equation 2) of linear extent L (see Reference [1] upon which this figure is based). Small patches lead to noisier disparity because of limited support in the intensity data. Larger patch sizes provide more support but produce displacement errors.

In order to model displacement errors, consider two locally constant surfaces separated by a discontinuity at $x = x'$. Assume that the original (or actual) surface property is:

$$d^*(x) = \delta d\, u(x' - x) = \begin{cases} \delta d & \text{for } x < x' \\ 0 & \text{for } x > x', \end{cases} \tag{3}$$

where δd is the magnitude of the discontinuity difference (sometimes referred to simply as 'the discontinuity') about the site x' and $u(x)$ is the Heaviside function. If we

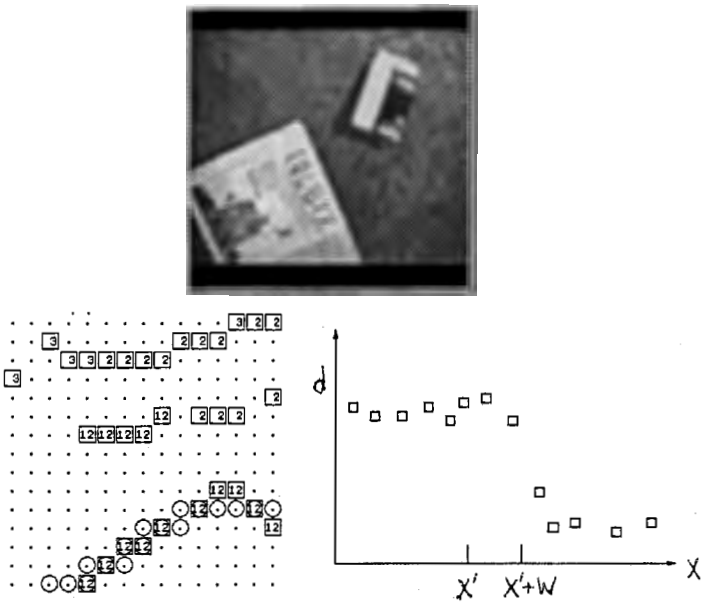

Figure 2: Displacement errors in stereo disparity data are illustrated for a stereo pair with the left image shown above. The lower left image shows sparse disparities (squares) and the object boundary (circles) for a small region of the scene. Disparities from the lower object (at disparity 12) are mistakenly displaced into the upper object (at disparity 2). These errors are characteristic of correlation-based early vision algorithms[1, 2]. The plot to the right is a slice of the sparse disparity data. The "indicated" discontinuity at $x' + W$ is displaced from the "actual" discontinuity at x'. Generally discontinuity detectors compute discontinuities at $x' + W$; however, an ideal detector should produce discontinuities at x'.

introduce a displacement error similar to those illustrated in Figures 2 and 3, the input to the discontinuity detector can be modeled as:

$$d(x) = \delta d \; u(x' + W - x) \tag{4}$$

where W is the *displacement shift* and it characterizes displacement errors. Our emphasis is on the factor W.[1] Consideration of the matching process and close examination of Figure 3 suggests that the displacement shift, W, is commensurate with L, the extent of the matching region. Consequently, W can be quite large; this

[1]Usually the discontinuity detector input is modeled as $d(x) = \delta d \; u(x' - x) + n(x)$ with $n(x)$ assumed to be spatially white, Gaussian noise. This model for $d(x)$ does not account for displacement errors. Extending our formulation in Equation 4 to include a Gaussian noise process presents no problem.

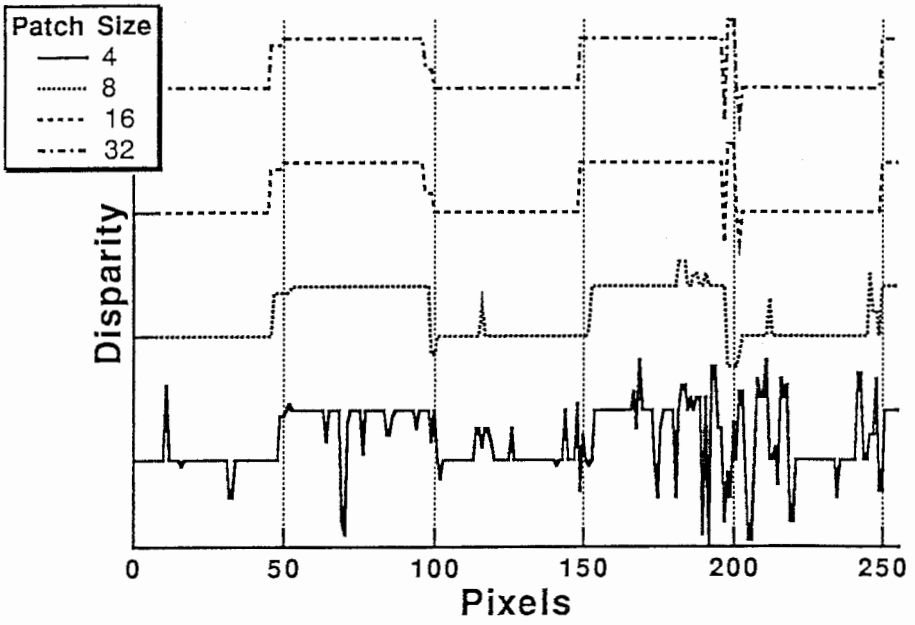

Figure 3: Stereo disparity as a function of patch size (based on Reference [1]). Discontinuities at 50, 100, 150, and 200 are introduced into synthetically created image data. A pixel-based stereo algorithm[2] computed the disparity. Near the introduced discontinuities, the computed disparities are systematically displaced from the actual discontinuity location.

is not a one or two pixel "registration-like" error. It is a significant distortion of the object boundary.

The preceding discussion has illustrated the nature of displacement errors. These displacement errors arise because of the multimodal behavior of the correlation function near discontinuities. The multimodal behavior arise as a consequence of violating the smoothness assumption used in early vision algorithms and because of the inherent uncertainly in occluded regions. The examples have illustrated displacement errors in stereo disparity; optical-flow algorithms commonly exhibit this type of error also.

The significance of these pervasive displacement errors has been all but ignored in surface reconstruction and discontinuity detection algorithms. A possible explanation as to why these errors have been overlooked is that most researchers in discontinuity detection are also interested in edge detection (e.g. Reference [5]). Edge detection is the process of finding discontinuities in intensity. However, because of the optics of image formation, intensity data is not plagued by displacement errors and as the previous description of displacement errors so clearly illustrate, the input noise in

edge and discontinuity detection must be considered fundamentally different

A simple question remains: how should a discontinuity detector (or surface reconstruction algorithm) handle these displacement errors? Based on the surface property data only, a discontinuity detector *must* mark discontinuities as displaced from their actual location. Yet, clearly this results in an error in the detected discontinuities and in the reconstructed surface. In fact, the surface property data alone leaves no signature to distinguish displacement errors. An additional source of information is required and hence integration of visual information must be performed.

3 Constrained Kernel Formulation

In this formulation, we intend to show that discontinuity detection requires a two stage process: first, smoothing to eliminate the displacement errors but while also preserving discontinuities and second, detecting discontinuities. The second stage is simply Canny's non-maximum suppression and hysteresis[4] applied to maxima in the gradient magnitude. Our interest here is in the smoothing stage in which the postulated discontinuities lead to locally variable kernels that can eliminate the displacement errors. These locally variable kernels are a consequence of solving the diffusion equation with local, static boundary conditions.

It is well known that solution to the heat equation,

$$\frac{\partial^2 f(x,t)}{\partial x^2} = \frac{\partial f(x,t)}{\partial t},$$

with diffusivity $c(x) = 1$ for a temporal impulse at $t = 0$ is[12]:

$$f(x,t) = \int_{-\infty}^{\infty} (\pi 4t)^{-\frac{1}{2}} e^{-(x-y)^2/4t} f^0(y)\, dy \tag{5}$$

where $f^0(x) \equiv f(x, 0^+)$. If at $x = x'$ we impose mirror boundary conditions, $f(x,t) = f(2x' - x, t)$, and seek the response at $4t = 2\sigma^2$ for $x > x'$ then

$$f_\sigma(x|x') = \int_{x'}^{\infty} \left[g_\sigma(y - (2x' - x)) + g_\sigma(y - x)\right] f^0(y)\, dy \tag{6}$$

where x' is the position of the boundary and $g_\sigma(x)$ is a Gaussian distribution with

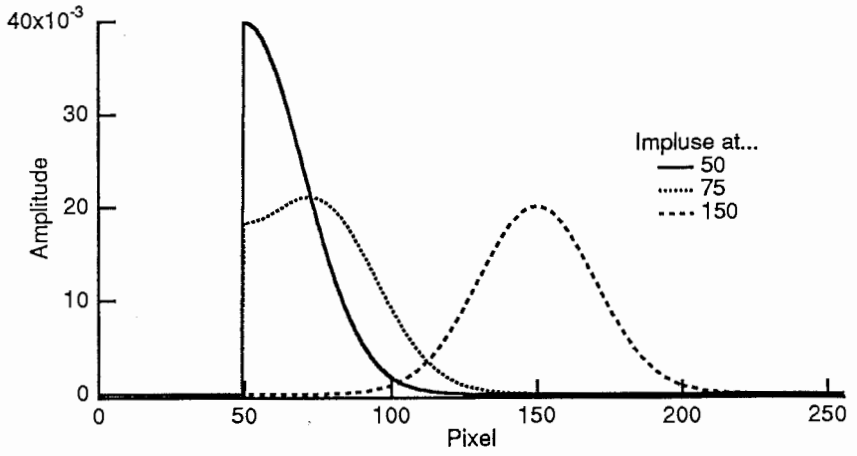

Figure 4: Three one-dimensional kernels ($\sigma = 20$ pixels) from Equation 7 for a postulated discontinuity at $x' = 50$ and impulses at 50, 75, and 150 pixels.

standard deviation σ. Because of the mirror boundary condition at x', this equation is seen simply as the response without boundary conditions to two impulses; one at $x = 0$ and its mirror impulse at $x = 2x'$. For a spatial impulse $f^0(y) = \delta(y)$ the Green's function is:

$$G_\sigma(x|x') = g_\sigma(x) + g_\sigma(2x' - x). \tag{7}$$

This Green's function is a *constrained kernel* for an impulse at $x = 0$ and a mirror boundary at $x = x'$. For several impulse locations and for a fixed σ, constrained kernels are illustrated in Figure 4.

The constrained kernels are themselves symmetric about the postulated discontinuity. Consequently if, over the spatial extent of the kernel, the input is symmetric about the postulated discontinuity, then the response will also be symmetric. For example, if a surface is locally constant about a postulated discontinuity, subsequent to smoothing, no discontinuity will be introduced. Such a situation arises when a discontinuity is incorrectly postulated; this is common and desired because postulated discontinuities, to be effective, should be a superset of the actual discontinuities. When intensity edges serve to postulate discontinuities, surface markings or texture will introduce extraneous postulates. However, these extraneous postulates will not introduce incorrect discontinuities because the response of Equation 6 is symmetric near the postulated discontinuity.

Subsequent to smoothing with the constrained kernels, discontinuities are identi-

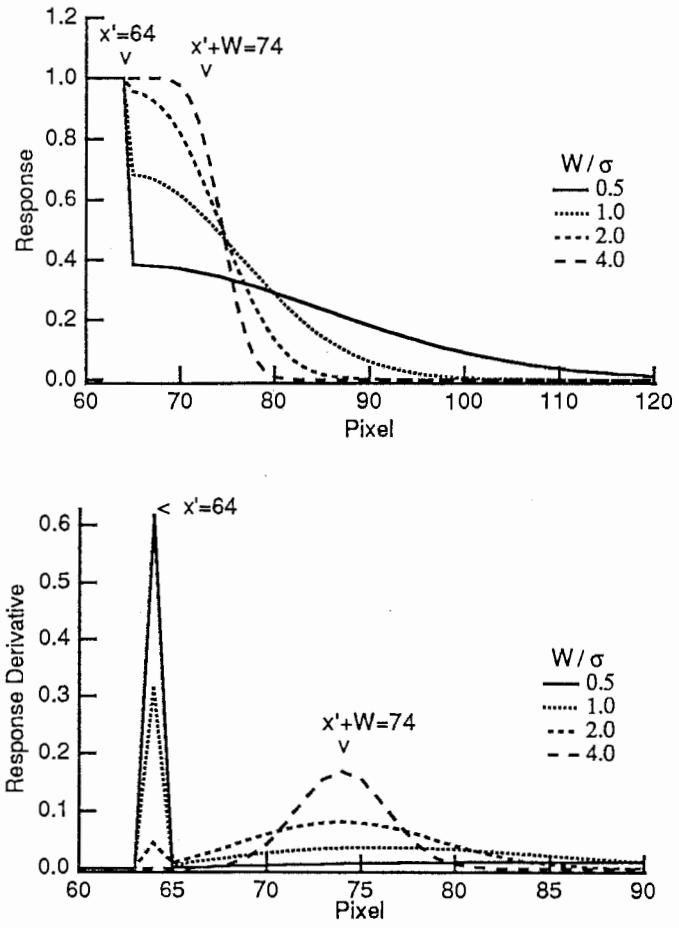

Figure 5: The response from Equation 6 is shown for different values of W/σ. The corrupted input is a unit step located at $x' + W = 74$ but the actual discontinuity is at $x' = 64$. A postulated discontinuity is assumed at x'. As σ increases ($W/\sigma \rightarrow 0$), the discontinuity is shifted to x' and thus the displacement error is eliminated. Also, for large values of σ and for $x > x'$, the surface smoothly decreases. Later, when applied to real images, this smooth decrease in the response for $x' < x < x' + W$ will highlight the pervasiveness of displacement errors and the success of constrained kernels in eliminating the errors. The lower plot illustrates the derivative of the response. When $W/\sigma > 2$ the derivative is peaked strongly at $x' + W$; yet, increasing σ suppressing this peak and shifts the response to x'. Note that the two plots have a different scale for the x axis.

fied as peaks in the gradient magnitude. Following Canny's hysteresis scheme: weak peaks are discarded, strong peaks are retained, and other peaks are retained if they belong to a contour containing a strong peak. Small discontinuities at a postulated discontinuity, whether due to noise or an actual discontinuity, remain and can be discarded based on a selected noise threshold.

We are still faced with the problem of eliminating displacement errors. Consider the disparity $d(x)$ of Equation 4 as the corrupted input for an actual surface given in Equation 3 as $d^*(x)$. The disparity $d(x)$ shows no discontinuity at x'; it has been displaced to $x' + W$. If we smooth $d(x)$ with the constrained kernels based on Equation 6, the response is:

$$f_{\{\frac{W}{\sigma},\delta d\}}(x|x') = \delta d \begin{cases} \int_0^{\frac{W}{\sigma}} [g_1(y + (x - x')) + g_1(y - (x - x'))]\, dy & \text{for } x > x' \\ 1 & \text{for } x < x' \end{cases} \tag{8}$$

A family of these responses are illustrated in Figure 5 for several values of W/σ. For small amounts of smoothing ($W/\sigma > 2$) the response remains large near $x' + W$. For this small σ the constrained kernel is essentially Gaussian at $x' + W$. The postulated discontinuity does not contribute significantly to the response at $x' + W$ and therefore in this case our detector fails, as do other discontinuity detectors, by marking a discontinuity incorrectly near $x' + W$. If, however, σ is increased ($W/\sigma < 2$) so that the postulated discontinuities play are role near $x' + W$; the constrained kernels deviate from a Gaussian and the discontinuity shifts from $x' + W$ back to x'. Figure 5 shows this effect very clearly as a function of decreasing W/σ.

This discontinuity that is introduced at $x = x'$ can be detected if it is above the noise threshold. The detectability places a restriction on W/σ. Define δd_{max} as the largest disparity difference at a displacement error and δd_{min} as the smallest disparity difference required to mark a discontinuity. (Typically, $\delta d_{min} \sim 1, 2$ pixels and δd_{max} is the full range of the disparity; maybe 20, 25 pixels.) Also, define δd_{th} as the disparity difference above which a discontinuity should be marked. To ensure that a discontinuity at x' is marked *and* that the displacement error at $x' + W$ is not marked, we have two conditions:

$$\begin{aligned} \left| f_{\{\delta d_{max}\}}(x' + W + \epsilon|x') - f_{\{\delta d_{max}\}}(x' + W - \epsilon|x') \right| &< \delta d_{th} \\ \left| f_{\{\delta d_{min}\}}(x' + \epsilon|x') - f_{\{\delta d_{min}\}}(x' - \epsilon|x') \right| &> \delta d_{th} \end{aligned} \tag{9}$$

where $\epsilon \to 0$ and the dependence of $f()$ on W/σ has been suppressed for display simplicity. The first equation states that after smoothing with the constrained kernel

even the largest displacement error of δd_{max} should not produce a disparity difference over the threshold. The second equation states that the discontinuity difference at $x = x'$ should be over the threshold so that the discontinuity is correctly marked at x'. Consideration of these equations and examination of Figure 5 indicate that increasing σ ensures that the above two conditions are met.

Discontinuities tend to be marked as a subset of the postulated discontinuities because of the shift induced by the constrained kernels. Another approach is to explicitly restrict the discontinuities to a subset of the postulated discontinuities; then the first equation in Equations 9 need not be satisfied and the condition on σ is not as strict. With sufficient smoothing to shift displacement errors both approaches perform well. The results in Section 5 uses the first approach where discontinuities could be marked anywhere.

4 Relation to other Discontinuity Detectors

Although our approach is related to several other approaches such as: anisotropic diffusion[13], adaptive windows[14], graduated non-convexity[15], computational molecules[16], and analog VLSI[17]; the use of postulated discontinuities is unique and, as Sections 2 and 3 have detailed, is mandated by the non-spatially white noise that pervades surface property data. These other approaches detect discontinuities admirably when applied to intensity images, range data, or synthetically created, white noise corrupted 'test' images because these applications do not exhibit displacement errors. As Figure 5 and the associated discussion illustrates, unless postulated discontinuities are used, these discontinuity detector will fail by marking discontinuities at the displaced position $x' + W$ rather than at the actual location x'.

Some early discontinuity detectors used postulated discontinuities for prior information[6, 7]. These early detectors were based on Markov random fields (Mrfs) and in hindsight suffered from needless complexity. In particular, Mrf formulations are computationally demanding and consequently numerous techniques have arisen to circumvent the computational aspects[18, 19]. In addition, Mrfs require specification of numerous parameters to impose smoothness and continuity on the solution. Specification of these parameters is difficult (e.g. Reference [20]). When postulated

discontinuities are used, smoothness of discontinuities arises naturally and the computational demands are nearly identical to those of repeated binomial convolution (as described in the next section).

Also, most other approaches seek simultaneously to compute the discontinuities and to reconstruct the surface property. One example is anisotropic diffusion where $c(x)$ can vary as the diffusion proceeds. The smoothed surface property data *alone* is used to compute $c(x)$; a large gradient in the surface property produces a small value for $c(x)$ thereby inhibiting diffusion (i.e. smoothing). Figure 5 showed that if $c(x)$ is allowed to vary then an early commitment to a discontinuity is possible and is likely to occur at its displaced position $x' + W$. As shown very clearly within the context of scale space for adaptive windows[14], once committed, the reconstructed surface will remain in error. Smoothing must be performed prior to discontinuity detection to allow the displacement errors to shift. This is the approach advocated here; the sequential process of smoothing with constrained kernels followed by detection of the discontinuities.

A related integration technique based on mean field theory[21] and a proposed technique to detect discontinuities in analog VLSI hardware[17] can be easily extended to compensate for displacement errors. Both techniques include concepts related to postulated discontinuities however, both simultaneously smooth the surface property and detect the discontinuities. If the postulated discontinuities are used as the initial state for the line process[22] and then then line process is held fixed while the surface is smoothed, the discontinuity can shift to its correct location at x'.

5 Discontinuity Detection Results

Discontinuity detection based on the constrained kernels from postulated discontinuities is highly efficient. This section briefly describes implementation of the constrained kernels and the extension to two-dimensions. Results for discontinuity detection from stereo disparity and optical-flow magnitude are presented as well.

The response at each pixel in the image is based on convolution with a unique kernel, the constrained kernel. Convolution with these locally variable constrained kernels would be computationally impractical were it not for binomial convolution.

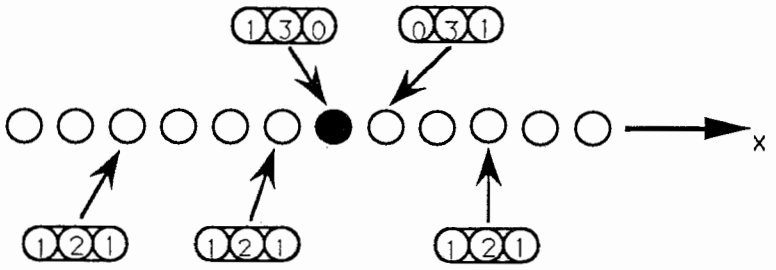

Figure 6: Binomial masks for various positions relative to a postulated discontinuity, shown with a filled circle, are illustrated. Smoothing between the postulated discontinuity and one of its neighbors is inhibited. The binomial masks in this case are given as $(1,3,0)$ and $(0,3,1)$. Away from the postulated discontinuity the mask is $(1,2,1)$. All masks are subsequently normalized to 1.

Repeated convolution with the binomial mask of Figure 6 produces an approximation to a Gaussian convolution with $\sigma = \sqrt{(1+N)/2}$ where N is the number of binomial iterations. Mirror boundary conditions are imposed by modifying the binomial mask near a postulated discontinuity. As Figure 6 illustrates, when one neighbor is a postulated discontinuity smoothing with that neighbor is inhibited. These masks can all be computed locally and are thus amenable to parallel computation. Our results were obtained with a 16k CM-2 Connection Machine[23].

The extension to two-dimensions is implemented by sequential application of binomial convolution along each of the two axes to build up the constrained kernels. Examples of 2D constrained kernels are shown in Figure 7. For these kernels and the subsequent results $\sigma = 20$ pixels. This large σ is mandated by the nature of displacement errors as described in Equation 9. The window size used for both the stereo and optical-flow algorithms is typically between 16 and 32 pixels and the magnitude of the disparity difference is on the order of 20 pixels.

Figure 8 shows two examples of discontinuity detection using the constrained kernels on stereo disparity data. The feature-based stereo algorithm[24] produces disparities ranging from 0 to 22 pixels. The patch size for the stereo correspondence was 16 x 16 pixels. The sparse disparities were converted to a dense field with simple region growing; this is labeled in Figure 8 (and 9) as "Filled Surface." The postulated discontinuities were based on intensity edges computed with Canny's edge detector.

The right side of Figure 8 shows the results from our detector. The discontinuities

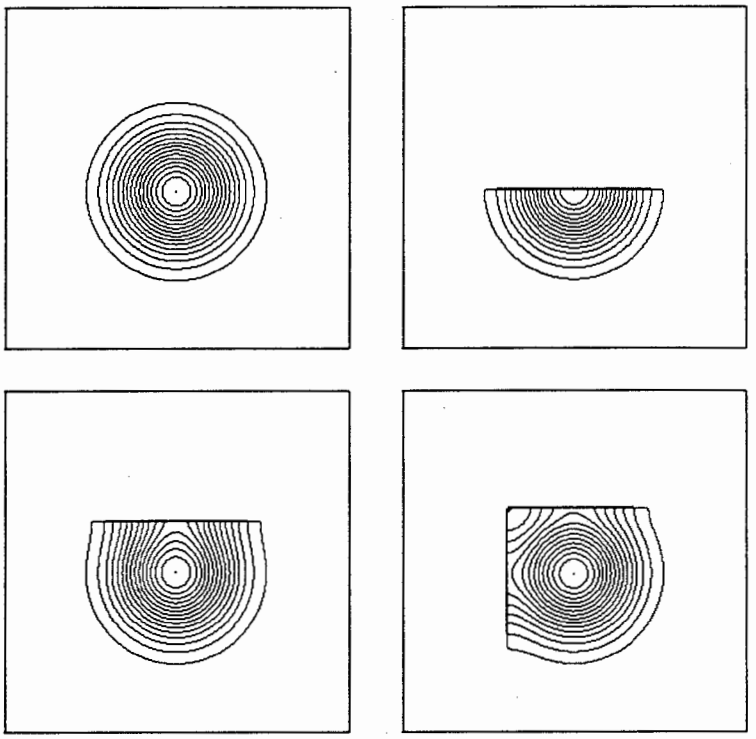

Figure 7: Contour plots of two-dimensional kernels generated with binomial convolution from an impulse at (128,128) with $\sigma = 20$. Clockwise from the upper left, the postulated discontinuities are: none, line at $x = 127$, line at $x = 85$, two lines at $x = 75$ and $y = 75$. Each plot is scaled differently; all kernels are normalized to 1.

are located accurately relative to the actual object boundaries (as observed visually). The smoothed disparity is also shown. Note that near the discontinuities the surface slowly decays away. This decay was described in Figure 5 and it clearly identifies those regions in the image plagued by displacement errors. As the figures illustrate, nearly all the discontinuities were computed accurately in spite of these errors.

Figure 9 shows similar results based on optical-flow[11]. The surface shown is the magnitude of the optical-flow. The discontinuities correspond to the boundaries between objects with different motions. The texture in regions of uniform motion have been eliminated. Once again the decay attests to the pervasiveness of the displacement errors and to the validity of our detection technique.

It is clear that the large value of σ needed to compensate for the displacement errors will oversmooth in regions far from the discontinuity. Local variation in σ is possible although our results have not exploited this option. A uniform σ has not

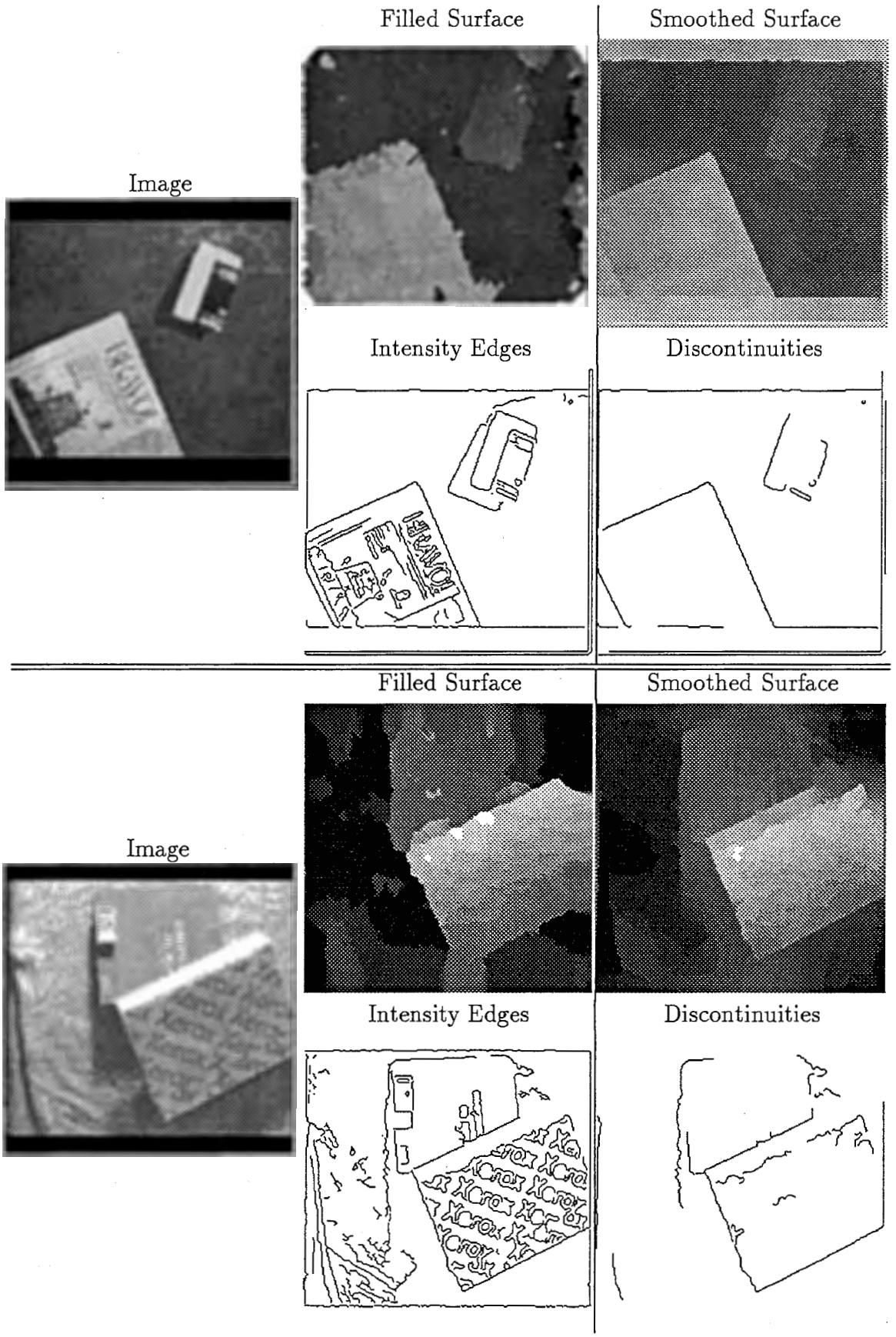

Figure 8: Depth discontinuities from stereo. Inputs on left, outputs on right. See page 15.

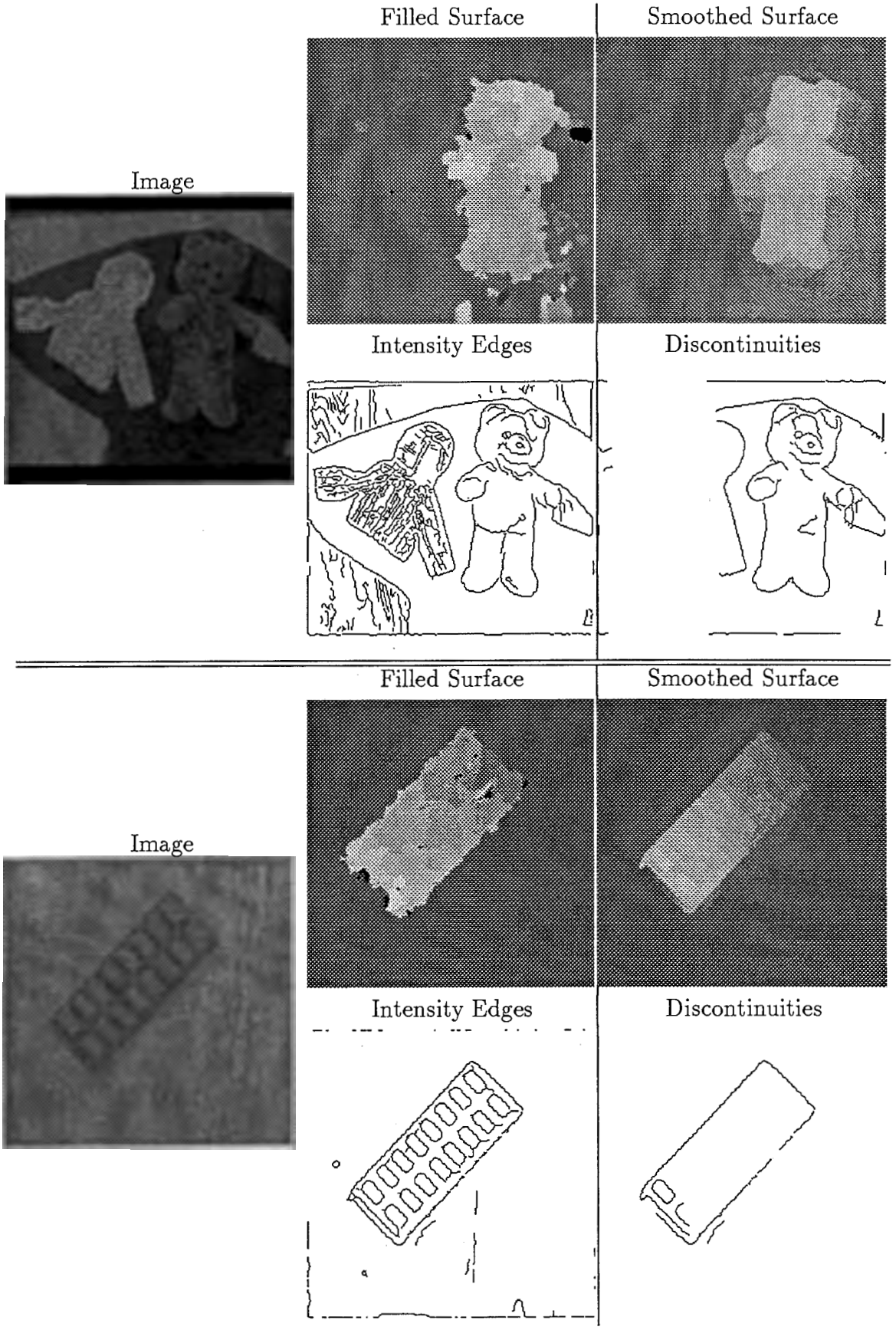

Figure 9: Discontinuities in optical-flow magnitude. Inputs on left, outputs on right. See page 15.

detracted from our results both in discontinuity detection and surface reconstruction.

6 Summary

Discontinuity detection differs from edge detection because of the noise in the surface property data. As we have shown, the pervasive noise in surface property data occurs at precisely those locations we seek to detect: the discontinuities. These so called displacement errors can only be resolved if prior information regarding their location is known. We have introduced the notion of postulated discontinuities to provide this prior information and have described a smoothing technique, based on constrained kernels, to compensate for the displacement errors.

The postulated discontinuities need not be intensity edges. In our work this has proved quite successful as the results in Section 5 illustrated. However, if other prior information more accurately identifies potential discontinuities then clearly intensity edges could be supplanted. Our use of intensity edges to postulate discontinuity locations rests upon the fundamental assumption of passive vision system: changes in surface properties originate intensity variations. Clearly this assumption fails in certain well-documented cases; second order texture boundaries and isoluminant color boundaries are two such cases. Similarly, for laser range data it is irrelevant to speak of intensity variations. However, for vision processing in a general environment, intensity variations arise and without intensity variations, low-level vision algorithms like stereo and motion will themselves fail. Hence the fundamental importance of intensity edges in computational vision.

Although our discontinuity detector was described in the context of surface reconstruction, the reliance on intensity edges suggests that the detector also could be characterized as a labeling or perceptual grouping algorithm. The detector associates the type of discontinuity (in the cases described herein: depth and motion, others are possible) with an intensity edge. These labels should facilitate higher-level processing such as feature grouping, segmentation, and recognition and the reduction in image clutter should improve the combinatorics of these higher-level algorithms.

References

[1] Masatoshi Okutomi and Takeo Kanade. A locally adaptive window for signal matching. Technical Report CMU-CS-90-178, Carnegie Mellon University, Pittsburgh, PA, October 1990.

[2] James J. Little and Walter E. Gillett. Direct evidence for occlusion in stereo and motion. *Image and Vision Computing*, 8(4):328–340, November 1990.

[3] E. B. Gamble Jr. *Integration of Early Vision Cues for Recognition.* PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, 1990.

[4] John F. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.

[5] David Lee. Coping with discontinuities in computer vision: Their detection, classification, and measurement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(4):321–344, April 1990.

[6] Edward B. Gamble and Tomaso Poggio. Visual integration and detection of discontinuities: The key role of intensity edges. A.I. Memo No. 970, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, October 1987.

[7] Paul B. Chou and Christopher M. Brown. Multi-modal segmentation using Markov random fields. In *Proceedings Image Understanding Workshop*, pages 663–670, Los Angeles, CA, February 1987. Morgan Kaufmann, San Mateo, CA.

[8] Tomaso Poggio, Edward B. Gamble, and James J. Little. Parallel integration of vision modules. *Science*, 242:436–440, 1988.

[9] Anselm Spoerri and Shimon Ullman. The early detection of motion boundaries. In *Proceedings of the International Conference on Computer Vision*, pages 209–218, London, England, June 1987. IEEE, Washington, DC.

[10] Tomaso Poggio, Vincent Torre, and Christof Koch. Computational vision and regularization theory. *Nature*, 317:314–319, 1985.

[11] Heinrich H. Bülthoff, James J. Little, and Tomaso Poggio. A parallel algorithm for real-time optical flow. *Nature*, 337:549 – 553, February 1989.

[12] Gilbert Strang. *Introduction to Applied Mathematics.* Wellesley-Cambridge Press, Wellesley, Massachusetts, 1986.

[13] Pietro Perona and Jitendra Malik. Scale space and edge detection using anisotropic diffusion. In *IEEE Computer Society Workshop on Computer Vision*, Miami, Florida, 1987.

[14] Philippe Saint-Marc, Jer-Sen Chen, and Gérard Medioni. Adaptive smoothing: A general tool for early vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6):514–529, June 1991.

[15] Andrew Blake and Andrew Zisserman. *Visual Reconstruction.* MIT Press, Cambridge, Mass, 1987.

[16] Demetri Terzopoulos. The role of constraints and discontinuities in visible-surface reconstruction. In *Proceedings IJCAI*, pages 1073–1077, August 1983.

[17] John G. Harris, Christof Koch, and Jin Luo. A Two-Dimensional analog VLSI circuit for detecting discontinuities in early vision. *Science*, 248:1209–1211, June 1990.

[18] Julian Besag. On the statistical analysis of dirty pictures. *J. Roy. Stat. Soc.*, B48:259–302, 1986.

[19] Paul B. Chou and Christopher M. Brown. Multimodal reconstruction and segmentation with Markov random fields and HCF optimization. In *Proceedings Image Understanding Workshop*, pages 214–221, Cambridge, MA, April 1988. Morgan Kaufmann, San Mateo, CA.

[20] Bernard W. Silverman, Christopher Jennison, Julian Stander, and Timothy C. Brown. The specification of edge penalties for regular and irregular image pixels. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(10):1017–1024, October 1990.

[21] Davi Geiger. *Visual Models with Statistical Field Theory.* PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, 1989.

[22] Stuart Geman and Don Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

[23] William Daniel Hillis. *The Connection Machine.* PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, 1985.

[24] Michael Drumheller. Connection Machine stereo matching. In *Proceedings AAAI*, pages 748–753, August 1986.