IENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 1988	memorandum

4. TITLE AND SUBTITLE

Analysis of Differential and Matching Methods for Optical Flow

5. FUNDING NUMBERS

DACA76-85-C-0010
N00014-85-K-0124

6. AUTHOR(S)

James J. Little and Alessandro Verri

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Artificial Intelligence Laboratory
545 Technology Square
Cambridge, Massachusetts 02139

8. PERFORMING ORGANIZATION REPORT NUMBER

AIM 1066

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
Information Systems
Arlington, Virginia 22217

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

None

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Distribution of this document is unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Abstract: Several algorithms for optical flow are studied theoretically and experimentally. Differential and matching methods are examined; these two methods have differing domains of application - differential methods are best when displacements in the image are small (< 2 pixels) while matching methods work well for moderate displacements but do not handle sub-pixel motions. Both types of optical flow algorithm can use either *local* or *global* constraints, such as spatial smoothness. Local matching and differential techniques and global differential techniques will be examined.
Most algorithms for optical flow utilize weak assumptions on the local variation of the flow and on the variation of image brightness. Strengthening these assumptions improves the flow computation. The computational consequence of this is a need for larger spatial and temporal support. Global differential approaches can be extended to local

(continued on back)

14. SUBJECT TERMS (key words)		15. NUMBER OF PAGES
optical flow	numerical differentiation	27
computational vision		16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED

(patchwise) differential methods and local differential methods using higher derivatives. Using larger support is valid when constraints on the local shape of the flow are satisfied. We show that a simple constraint on the local shape of the optical flow, that there is slow spatial variation in the image plane, is often satisfied. We show how local differential methods imply the constraints for related methods using higher derivatives. Experiments show the behavior of these optical flow methods on velocity fields which do not obey the assumptions. Implementation of these methods highlights the importance of numerical differentiation. Numerical approximation of derivatives requires care, in two respects: first, it is important that the temporal and spatial derivatives be matched, because of the significant scale differences in space and time, and, second, the derivative estimates improve with larger support.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 1066 August 1988

ANALYSIS OF DIFFERENTIAL AND MATCHING METHODS FOR OPTICAL FLOW

James J. Little and Alessandro Verri

Abstract: Several algorithms for optical flow are studied theoretically and experimentally. Differential and matching methods are examined; these two methods have differing domains of application - differential methods are best when displacements in the image are small (< 2 pixels) while matching methods work well for moderate displacements but do not handle sub-pixel motions. Both types of optical flow algorithm can use either *local* or *global* constraints, such as spatial smoothness. Local matching and differential techniques and global differential techniques will be examined.
Most algorithms for optical flow utilize weak assumptions on the local variation of the flow and on the variation of image brightness. Strengthening these assumptions improves the flow computation. The computational consequence of this is a need for larger spatial and temporal support. Global differential approaches can be extended to local (patchwise) differential methods and local differential methods using higher derivatives. Using larger support is valid when constraints on the local shape of the flow are satisfied. We show that a simple constraint on the local shape of the optical flow, that there is slow spatial variation in the image plane, is often satisfied. We show how local differential methods imply the constraints for related methods using higher derivatives. Experiments show the behavior of these optical flow methods on velocity fields which do not obey the assumptions. Implementation of these methods highlights the importance of numerical differentiation. Numerical approximation of derivatives requires care, in two respects: first, it is important that the temporal and spatial derivatives be matched, because of the significant scale differences in space and time, and, second, the derivative estimates improve with larger support.

Acknowledgments. This report describes research done within the Artificial Intelligence Laboratory. Support for the A.I. Laboratory's artificial intelligence research is provided by the Advanced Research Projects Agency of the Department of Defense under Army contract DACA76-85-C-0010 and in part under Office of Naval Research (ONR) contract N00014-85-K-0124. Support for this research is also provided by a grant from ONR, Engineering Psychology Division, and by a Hughes Aircraft Corporation gift to the Artificial Intelligence Center for T. Poggio.

1 Introduction

Useful visual information can be extracted from optical flow — a $2D$ vector field which estimates the velocity of points in space as projected on the image plane. For example, images can be segmented at discontinuities of optical flow, while $3D$ motion and structure parameters can be computed from flow features in several ways [Wax84, Kan85, DN82b, DN82a, LHP80, TH84, WU85, BH83, WN86, Ull83]. A thorough analysis of how to recover optical flow robustly from a sequence of time-varying images, therefore, is important for understanding the strengths and weaknesses of the different methods that can be used.

In this paper, a number of algorithms for optical flow are examined on both theoretical and experimental grounds. In particular, we examine differential and matching methods. Emphasis has been given to algorithms which have a very simple and local computational structure. We contrast algorithms where optical flow is derived from simple local processing with those having global constraints. The former methods may require some simple smoothing to achieve coherence while the latter require global smoothing under constraint (regularization)[BPT88, YG88].

The flow can be derived either by *differential* or *matching* methods. Differential methods estimate image velocity from spatial and temporal variation in image brightness. Matching methods search for displacements which bring image brightness features into correspondence.

Our analysis suggests several points. The traditional algorithms for optical flow utilize weak assumptions on the local variation of the flow, and on the variation of image brightness. Strengthening these assumptions makes local flow computation possible. The computational consequence of stronger assumptions is a need for larger spatial and temporal support. Using larger support is valid when constraints on the local shape of the flow are satisfied. We show that a simple constraint on the local shape of the optical flow, slow spatial variation across the image plane, is often satisfied.

The initial measurements in differential methods include various spatial and temporal derivatives of the image. Much care must then be taken in numerical implementation of differentiation. For example, if the scale of the smoothing filter applied before differentiation is comparable with the size of the inter-pixel distance (σ is approximately 1) — which is usually the case — accurate numerical approximation of derivatives requires large spatial support. Intuitively, this happens because the inter-pixel distance is not "small" with respect to the spatial and temporal change in image brightness. Moreover, much care must be taken in mixing numerical approximations with different support. This fact plays a crucial role in analyzing performance of algorithms.

Rather general conclusions can be drawn from our analysis. Firstly, local algo-

rithms provide good measurements of optical flow and thus are particularly promising as inputs to later visual tasks. Moreover, suitable differential techniques are able to produce local estimates of the optical flow at much less computational expense than global methods.

Secondly, the use of simple constraints on the local shape of optical flow improves the quality of results. Localization and segmentation require precise estimates of the optical flow, while robustness demands an assumption of smoothness or coherence in the flow. These goals often conflict. The better the estimates in the initial stage, the less stringent need be the constraints applied in the smoothing stage.

The rest of the paper is organized in three sections. Section 2 discusses the general structure of methods for optical flow. In section 3 different algorithms are considered. The computational assumptions which underlie each of them are discussed and some novel theoretical arguments introduced. Experimental results are always presented to corroborate major points. Section 4 examines the performance of some of the algorithms when their assumptions do not hold. The concluding section summarizes our results.

2 General Structure of Methods for Optical Flow

This section discusses the general structure of several rather different algorithms which attempt to recover optical flow from a sequence of time-varying images.

The extraction of motion information from a sequence of images is a difficult task. Goals and computational constraints often do not match, and sometimes are conflicting. For example, local optical flow estimates can be refined by assuming different degrees of spatial smoothness. On the other hand, discontinuity of optical flow, localization of features, and $3D$ parameter estimates require accurate reconstruction of the optical flow. It proves useful to think of optical flow recovery as a process consisting of two steps. In the first step, *measurement*, the optical flow is estimated by means of local techniques. The measurements are not necessarily very accurate but should allow for a complete reconstruction of the optical flow, which is carried out in the second step, usually a *smoothing* or regularization step.

Indeed, traditional motion studies usually start from the assumption that the first step has to be incomplete and ineffective since the recovery of the optical flow is an underconstrained problem — the so-called "aperture problem" [MU81]. Recently, however, this assumption has been successfully challenged [UGVT88a, LBP88, RSE88]; several studies have proved experimentally and theoretically that there is enough information in a sequence of images to give an accurate measurement of local motion in a single step. A very important consequence of this result is the fact that often no further processing is required, particularly so when the measurement stage depends upon large spatial support, as in matching methods. In the next section, we will

examine in some detail the assumptions underlying these studies, and the structure and computational requirements of the algorithms that arise naturally from them. An influential algorithm proposed by Horn and Schunck[HS81] is also considered as a prototype of a certain way to look at the problem of computing the optical flow.

3 Algorithms for Optical Flow

In this section we review some algorithms for optical flow. Emphasis is given to computational issues and theoretical arguments which support the use of simple but effective constraints to produce measurements of optical flow on a strictly local basis.

3.1 Differential Algorithms

Optical flow information can be extracted by making assumptions about spatial and temporal variations of the image brightness. If $E = E(x_1, x_2, t)$ is the image brightness at the location $\vec{x} = (x_1, x_2)$ on the image plane at time t in a suitable system of orthogonal coordinates, the weakest possible assumption is that the total temporal derivative of the image brightness vanishes [FT79, HS81]:

$$\frac{dE}{dt} = 0 \tag{1}$$

Eq. 1 embodies the "brightness constancy" assumption. Since this is a very simple assumption, making no hypothesis on the spatial variation of the optical flow, it is the starting point of several methods for computing optical flow. Equation 1 can be read as an analytical formulation of the "aperture problem" — because from Eq. 1 it is only possible to recover the component of the optical flow in the direction of the spatial gradient. Equation 1 can be expanded into the total derivative of the optical flow $\vec{v} = (v_1, v_2)$:

$$\frac{\partial E}{\partial x_1} v_1 + \frac{\partial E}{\partial x_2} v_2 + \frac{\partial E}{\partial t} = 0 \tag{2}$$

Note that Eq. 2 involves only first spatial and temporal derivatives of the image brightness. Methods for optical flow based upon Eq. 2 need further constraints in order to determine the correct flow field. Horn and Schunck [HS81] introduced a smoothness assumption on the spatial variation of the optical flow, and chose the smoothest vector field which satisfies Eq. 2 in order to recover a unique solution. They also proposed an iterative algorithm to compute the solution, which we have implemented and tested. The update rule in the iterative scheme is

$$v_1{}^{n+1} = \bar{v_1}{}^n - \frac{\lambda(\frac{\partial E}{\partial x_1}\bar{v_1}{}^n + \frac{\partial E}{\partial x_2}\bar{v_2}{}^n + \frac{\partial E}{\partial t})}{1 + \lambda(\frac{\partial E}{\partial x_1}{}^2 + \frac{\partial E}{\partial x_2}{}^2)} \frac{\partial E}{\partial x_1}$$

$$v_2{}^{n+1} = \bar{v_2}{}^n - \frac{\lambda(\frac{\partial E}{\partial x_1}\bar{v_1}{}^n + \frac{\partial E}{\partial x_2}\bar{v_2}{}^n + \frac{\partial E}{\partial t})}{1 + \lambda(\frac{\partial E}{\partial x_1}{}^2 + \frac{\partial E}{\partial x_2}{}^2)} \frac{\partial E}{\partial x_2} \quad (3)$$

where superscripts denote the step and $\bar{v_1}, \bar{v_2}$ are local averages of v_1,v_2.[1] Figure 1 shows a "plaid" pattern, a superposition of vertical and horizontal sine waves of the same period (64 pixels) in an image of 256 × 256 pixels. The pattern is translating across the image plane from the upper left corner to the lower right corner by 6.4 pixels per frame. The Horn and Schunck algorithm recovers the correct translational flow for this pattern (Fig. 2), but requires many iterations of the iterative regularization (smoothing) step. Consider the effect of setting λ to zero in Eq. 3; then the regularization (as constructed) will choose the smoothest velocity field derived from repeated averaging of the initial measurements. The number of iterations of such a process is proportional to the square of the spatial scale, since it is a diffusion process. The spatial wavelength is large in this case (64 pixels), necessitating so many iterations to propagate information from regions where the image gradient is parallel to the flow to regions where the image gradient is normal to the flow (see Fig. 3 for the initial normal flow). When λ is not zero, the convergence is even slower, since the subtracted terms in Eq. 3 force the value of $\vec{v}$ away from the average value $(\bar{v_1}, \bar{v_2})$ toward the image gradient. Multigrid methods [Ter86] can be used to overcome some of these problems of scale.

To improve the measurement step without enforcing global smoothness constraints, a simple working assumption (examined later in Section 4) is that the optical flow does not change appreciably at neighboring pixels. Then, in the neighborhood of every pixel, Eq. 2 can be rewritten for the same unknown velocity. At each pixel, a linear system, possibly overconstrained, has to be solved (see [LK81, KTB87], among others). Every equation of the linear system for the optical flow $\vec{v} = (v_1, v_2)$ is of the form of Eq. 2, and v_1 and v_2 are assumed to be constant over a neighborhood of every pixel. This *local constraint* algorithm, applied to the sequence of Fig. 1, using a neighborhood of 11x11 pixels, produces the correct translational flow. Note that no further smoothing step has been implemented but that a larger support has been used. At least two questions arise from the implementation of this simple algorithm: firstly, what can be said about the assumption on the local structure of the optical flow, namely that the optical flow is locally "approximately constant"? An argument to support this and other similar constraints is presented in Section 4. Secondly, does this formulation lead to a seriously ill-conditioned algorithm? The performance of the algorithm is dependent on the local variation of the image gradient; the neighborhood must be large enough to contain sufficient variation of ∇E.

Other constraints, instead of Eq. 2, can be applied for optical flow. The total temporal derivative of several other quantities vanishes depending on the kind of

[1] This formula (due to Horn) corrects an error in the update rule given in [Hor85], p. 288.

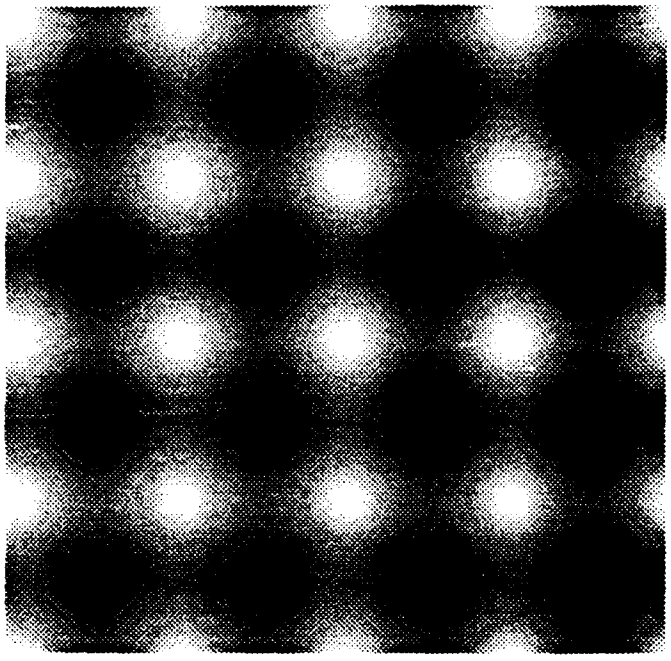

Figure 1: A "plaid" pattern, a superposition of vertical and horizontal sine waves of the same period (64 pixels). The pattern is translating across the image plane from the upper left corner to the lower right corner by 6.4 pixels per frame.

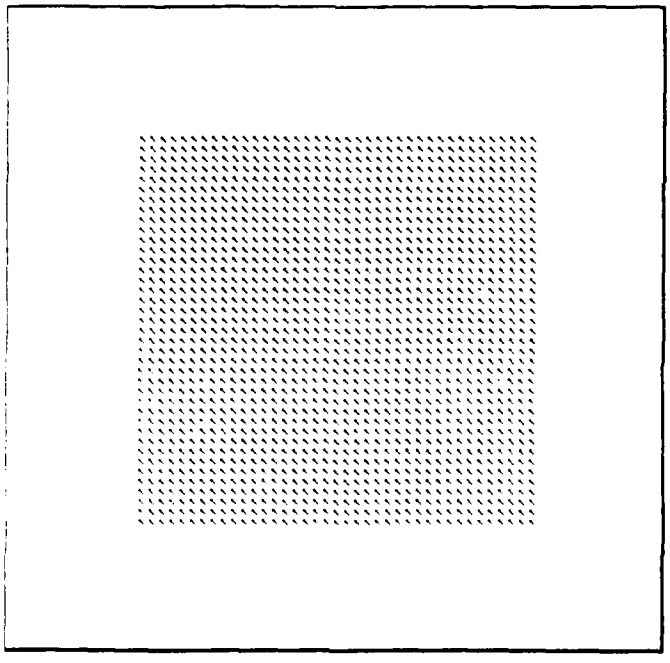

Figure 2: The optical flow field produced by the matching algorithm using a displacement range of ± 8 pixels; for this input, all algorithms performed similarly.

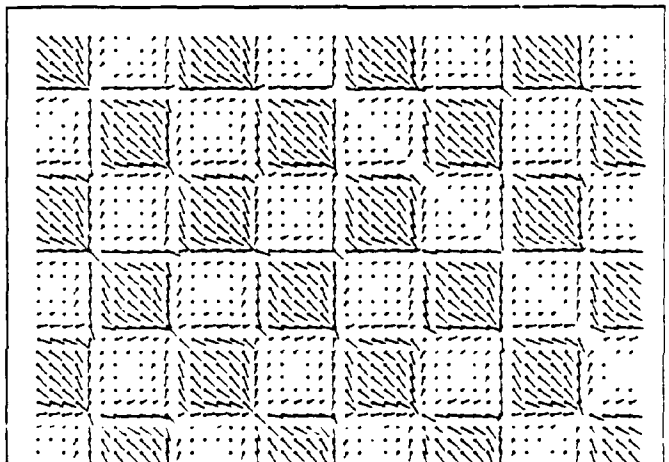

Figure 3: Normal flow obtained by solving Eq. 2 at the third of five frames. Image brightness derivatives have been computed by using five-point approximation of derivatives derived by means of Taylor's expansion.

observed motion [VGT90]. For example, the total temporal derivative of the spatial gradient vanishes for parallel translation on the image plane [UGVT88b], that is,

$$\frac{d\nabla E}{dt} = 0 \tag{4}$$

We term this assumption *gradient constancy*. Expanding Eq. 4 leads to the two constraints:

$$\begin{aligned} \frac{\partial^2 E}{\partial {x_1}^2} v_1 + \frac{\partial^2 E}{\partial x_1 \partial x_2} v_2 + \frac{\partial^2 E}{\partial x_1 \partial t} &= 0 \\ \frac{\partial^2 E}{\partial x_2 \partial x_1} v_1 + \frac{\partial^2 E}{\partial {x_2}^2} v_2 + \frac{\partial^2 E}{\partial x_2 \partial t} &= 0 \end{aligned} \tag{5}$$

at each pixel in the image. If the spatial changes in the optical flow can be assumed to be negligible (see Sec. 4) then Eq. 4 can be solved for the optical flow. However, since Eq. 4 provides two constraints for the optical flow, in principle the flow it produces needs less smoothing. Again this has been obtained by using larger support since it requires computing second derivatives [Nag83a, Nag83b, HL83, TP82, Nag87]. At locations where the two scalar equations in Eq. 4 are linearly dependent — that is, where the determinant of the Hessian, the matrix of second derivatives of image brightness, vanishes — there is not enough local information to measure the optical flow. Where the determinant of the Hessian is small, the linear system is often ill-conditioned. For an analysis of errors in local constraint methods, see [KTB87]. The gradient constancy algorithm performs well on the sequence in Fig. 1; where the Hessian vanishes, there are gaps - this implies the need of a small amount of smoothing to fill in gaps. Interestingly, the local constraint method and gradient constancy are intimately related. It is possible to show that the constraints given by locally solving the brightness constancy equation at three nearby points, assuming constant flow, imply the gradient constancy constraint.

Consider three points in the image, a central point p_1 and two others, p_2, displaced by $(\Delta x_1, 0)$ from p_1, and p_3, displaced by $(0, \Delta x_2)$ from p_1. The spatial and temporal derivatives of brightness at p_2 can be expanded in a Taylor series around p_1, letting superscripts denote the point at which derivatives are taken, to give

$$\begin{aligned} \frac{\partial E^2}{\partial x_1} &= \frac{\partial E^1}{\partial x_1} + \frac{\partial^2 E^1}{\partial {x_1}^2} \Delta x_1 \\ \frac{\partial E^2}{\partial x_2} &= \frac{\partial E^1}{\partial x_2} + \frac{\partial^2 E^1}{\partial x_2 \partial x_1} \Delta x_1 \\ \frac{\partial E^2}{\partial t} &= \frac{\partial E^1}{\partial t} + \frac{\partial^2 E^1}{\partial t \partial x_1} \Delta x_1 \end{aligned} \tag{6}$$

ignoring high order terms. Substituting these terms in Eq. 2 at p_2 yields

$$(\frac{\partial E^1}{\partial x_1} + \frac{\partial^2 E^1}{\partial {x_1}^2} \Delta x_1) v_1 + (\frac{\partial E^1}{\partial x_2} + \frac{\partial^2 E^1}{\partial x_2 \partial x_1} \Delta x_1) v_2 + (\frac{\partial E^1}{\partial t} + \frac{\partial^2 E^1}{\partial t \partial x_1} \Delta x_1) = 0. \tag{7}$$

Rearranging gives

$$(\frac{\partial E^1}{\partial x_1} + \frac{\partial E^1}{\partial x_2} + \frac{\partial E^1}{\partial t}) + (\frac{\partial^2 E}{\partial {x_1}^2} v_1 + \frac{\partial^2 E^1}{\partial x_2 \partial x_1} v_2 + \frac{\partial^2 E^1}{\partial t \partial x_1}) \Delta x_1 = 0. \quad (8)$$

The first sum vanishes, and, after division by Δx_1, we get the first equation in Eq. 5. Similar expansion at p_3 using second derivatives in x_2, with the same manipulations, yields the second equation in Eq. 5. Local solution of the brightness constraint equation at three points, chosen in this way, thus implies the gradient constancy constraint.

3.2 Matching Algorithms

Let us consider methods which match features from two images to derive displacements describing the optical flow field. The large class of correlation-based algorithms for stereo or motion are matching methods. [Dev75, Nis84, Ana87, LOY73, MP76, KMJ77] (see [Hor85] for references). As a prototype, we will use the parallel optical flow algorithm described in [LBP88]. This algorithm assumes that the optical flow is locally constant, that, for each point, the displacement of nearby points under the optical flow is the same as the displacement of the central point. We show later (Section 4) that this assumption is true at many points in most optical flow fields, particularly since the matching method only considers displacements at multiples of pixels. A second assumption states that the effects of foreshortening are small: note that this assumption is not true at occluding boundaries of rotating objects. In stereo, this assumption is more often violated, since the viewing angle (the angle between the projection ray of a surface point and the surface normal) varies much more between the two images.

The structure of the algorithm is as follows. At each point in the image, under each integer displacement, the image in the two frames are compared and a measure of the matching between points is computed, and summed over a small region. This can be interpreted as matching small patches from the first image with small patches in the second. Coherence of the resulting flow field is achieved as a result of the fact that the support regions, the patches, have large overlap. The displacement is chosen to maximize the matching measure over all displacements. There are two important parameters in this algorithm: the size of the summation regions and the range of displacements to be considered. For a further discussion of this algorithm and parameter choices, see [LB88].

The algorithm is simple and local, but is more computationally intensive than the measurement stage of most differential methods, since the number of comparison steps depends on the displacement range. The optical flow obtained by using this method on the sequence of images previously considered is shown in Fig. 2. We will see from later examples of the matching algorithm that the algorithm is robust under small deviations from the assumption of local constancy of the flow. When

the flow results from parallel translation in the image plane, and the magnitude of the displacement is integral and within the range of displacements considered, the matching algorithm should be exact, and thus can serve as a reference in some of our examples. Note that the gradient constancy method also produces correct estimates under these circumstances. The matching measure used in experiments reported here is the squared difference of the smoothed brightness over patches from the first and second images. Valid measurements result at locations where the image Hessian does not vanish. The constraint of gradient constancy leads to a matching criterion for the matching method, the squared difference of the image gradients, that can also be effectively used in the matching algorithm.

4 Verifying Assumptions

In this section we deal with computational and numerical assumptions made to recover the optical flow. Sequences of a rotating circle and of a planar surface which translates toward the observer are considered. Finally, the numerical implementation of derivatives for optical flow is discussed in detail.

4.1 Local Properties of the Optical Flow

Let us consider the simple case in which a planar surface is translating in space. The optical flow $\vec{v} = (v_1, v_2)$ can then be written as

$$\vec{v} = \frac{T_3}{f\gamma}\vec{\alpha} \cdot \vec{x}(\vec{x} - \vec{x}^s) \tag{9}$$

where $\vec{x} = (x_1, x_2, f)$ is the perspective projection onto the image plane of the point $\vec{X}$ in a system of coordinates defined by the triple of mutually orthogonal unit vectors $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$. The center of projection lies at the origin and the image plane has equation $X_3 = \vec{X} \cdot \vec{e}_3 = f$. The point $\vec{x}^s = f\vec{T}/T_3$, $T_3 = \vec{T} \cdot \vec{e}_3$, is the singular point of the optical flow, or the point where the optical flow vanishes. The vector $\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ is the unit normal to the planar surface which has equation

$$\gamma = \vec{\alpha} \cdot \vec{X} \tag{10}$$

Note that $\vec{x}$ is now considered a $3D$ vector. Its third component is, however, always equal to the focal length f. We want to estimate the spatial variation of the optical flow across the image plane. Let us define δv_{ij}, the relative spatial variation of v_i in the direction $\vec{e}_j$, as follows:

$$\delta v_{ij} = \frac{\left|\frac{\partial v_i}{\partial x_j}\right|}{\|\vec{v}\|} \tag{11}$$

where $\frac{\partial v_i}{\partial x_j}$ is the spatial partial derivative of v_i, $i = 1, 2$, in the direction $\vec{e}_j$, $j = 1, 2$, and $\|\vec{v}\|$ is the magnitude of the optical flow at $\vec{x}$. The relative spatial variations are good measures of the magnitude of spatial variation of the optical flow, because they do not depend on the magnitude of the flow field. Previous examinations of the uniformity of optical flow [KTB87] formulate a similar measure of flow variation and derive an approximate bound for the flow induced by a planar surface, although under more restricted assumptions on viewing geometry.

Now, straightforward computation of the partial derivatives of the v_i leads to the following inequality

$$\left|\frac{\partial v_i}{\partial x_j}\right| \leq \frac{T_3}{f|\gamma|}(|\vec{\alpha} \cdot \vec{x}| + \|\vec{x} - \vec{x}^s\|) \tag{12}$$

while for the magnitude of the optical flow we have

$$\|\vec{v}\| = \frac{|\vec{\alpha} \cdot \vec{x}| \cdot \|\vec{x} - \vec{x}^s\|}{f|\gamma|} \tag{13}$$

Thus, since $\|\vec{x}\| \geq f$, one obtains

$$\delta v_{ij} \leq \frac{1}{\|\vec{x} - \vec{x}^s\|} + \frac{1}{f \cos\theta} \tag{14}$$

where θ is the angle between the normal vector $\vec{\alpha}$ and $\vec{x}$. The focal length f is typically 10^3 pixels. Therefore, when the angle is smaller than 20 degrees, at locations thirty pixels away from the singular point the relative spatial variations are less than 5% percent per pixel. For angles as wide as 85 degrees, at least ten pixels away from the singular point the relative spatial variation is still less than about 10% per pixel.

A similar argument for pure rotation leads to the following inequality

$$\delta v_{ij} \leq \frac{2\|\vec{\omega}\|}{|\omega_3|\|\vec{x} - \vec{x}^s\|\sqrt{1 - 2\cos^2\theta}} \tag{15}$$

where $\vec{x}^s = f\vec{\omega}/\omega_3$, $\omega_3 = \vec{\omega} \cdot \vec{e}_3$ is the singular point of the optical flow and θ is the angle between $\vec{x} - \vec{x}^s$ and $\vec{x}$. In the case in which $\vec{\omega} = \omega\vec{e}_3$ (when the rotational and optical axis coincide), for example, Eq. 15 can be rewritten as

$$\delta v_{ij} \leq \frac{2}{\|\vec{x} - \vec{x}^s\|} \tag{16}$$

which is similar to Eq. 14. For $\omega_3 = 0$ or $|\omega_3| << \|\omega\|$, a similar conclusion can be obtained.

The above analysis can easily be extended to an arbitrary smooth surface, due to its local nature. The only difference is that, since the normal to the surface is changing, $\vec{\alpha}$ must change accordingly. As intuitively expected, therefore, locations of strong spatial variation lie near occluding boundaries (that is, where the normal to

the surface is almost perpendicular to the projecting ray) and probably not far away from discontinuities. Horn and Schunck[HS81] show that (in the case of rotational motion)

$$\nabla^2 v_1 = \omega_2 \nabla^2 x_3, \nabla^2 v_2 = \omega_1 \nabla^2 x_3, \tag{17}$$

showing that "the smoothness of the optical flow is directly related to the smoothness of a rotating body".

The above arguments suggest that the local shape of the optical flow is indeed fairly simple. Due to discretization errors and noise, at most locations, optical flows are likely to be locally indistinguishable from constant vector fields. Therefore, assumptions like local constancy, i.e., slow variation of the optical flow across the image plane, capture a fundamental local property of optical flow, entirely due to the simple structure of $3D$ rigid motion which has been, as usual, implicitly assumed.

4.2 Non-constant Velocity Fields

We consider the behavior of the algorithms when the assumptions of small spatial variation of the velocity are violated, for example, in the cases of rotational and looming fields.

4.2.1 Rotational Field

First we examine a sequence of images of a circle of random grey levels rotating on the image plane around its center at a fixed angular velocity of 3 degrees per frame, resulting in maximum displacement of 6.7 pixels. A random grey level circle rotates before a random grey level background parallel to the image plane. The radius of the circle is 128 pixels, random grey levels are uniformly distributed in the range 0..255, and the dots are 2 by 2 pixels, finally smoothed by a Gaussian, $\sigma = 2$. The output of the Horn-Schunck algorithm is shown in Fig. 5; large values of λ, the regularization parameter, select smoother optical flows. The results of the local constraint algorithm, and the gradient constancy algorithm (Eq. 4) are shown respectively in Fig. 6 and Fig. 7. The result of the matching algorithm is shown in Fig. 8. Both the local constraint algorithm and the gradient constancy algorithm are followed by a smoothing step to restore coherence. Convolving each of the two components of the optical flow resulting from the measurement step with a $2D$ symmetric Gaussian function produces more coherent optical flows. Note that the theoretical argument of the previous section, that small relative variation in the optical flow can be assumed, is clearly verified.

We have also compared the optical flow fields generated by the various algorithms with the true projected velocity field, which is available since these are synthetic images. Differential algorithms produce real-valued flow fields, while the matching

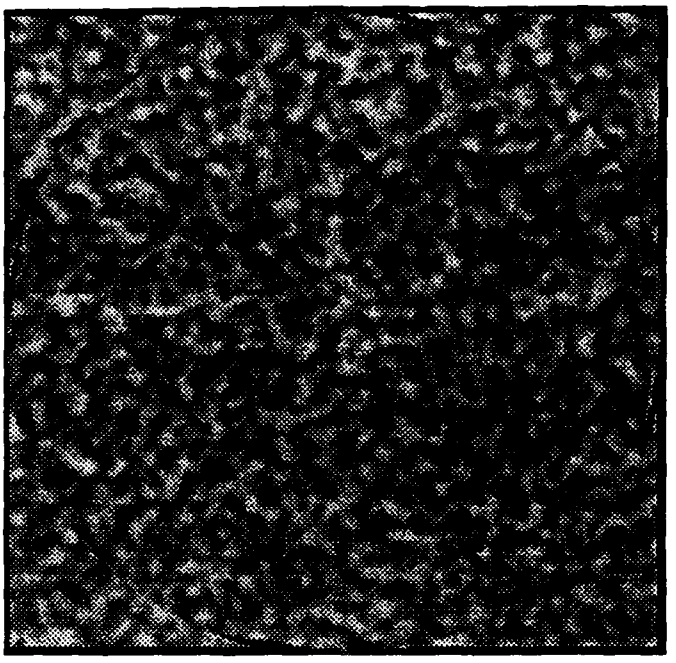

Figure 4: A random grey level disc rotates on a random grey level background parallel to the image plane. The disc radius is 128 pixels; grey levels are uniformly distributed from 0..255; the dots are 2 by 2 pixels, smoothed by a Gaussian, $\sigma = 2$.

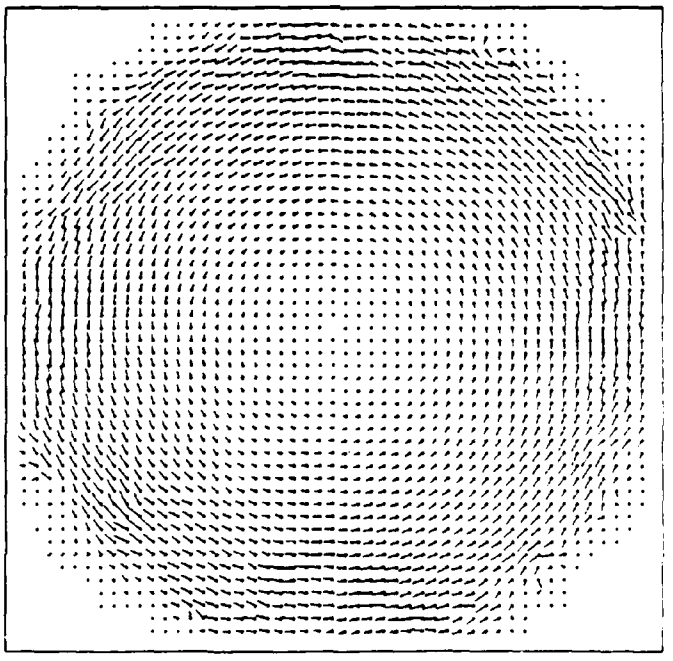

Figure 5: The optical flow field (rotating circle) produced by the Horn-Schunck algorithm, as in [Hor85], after 400 iterations of a numerical algorithm with $\lambda = 1$.

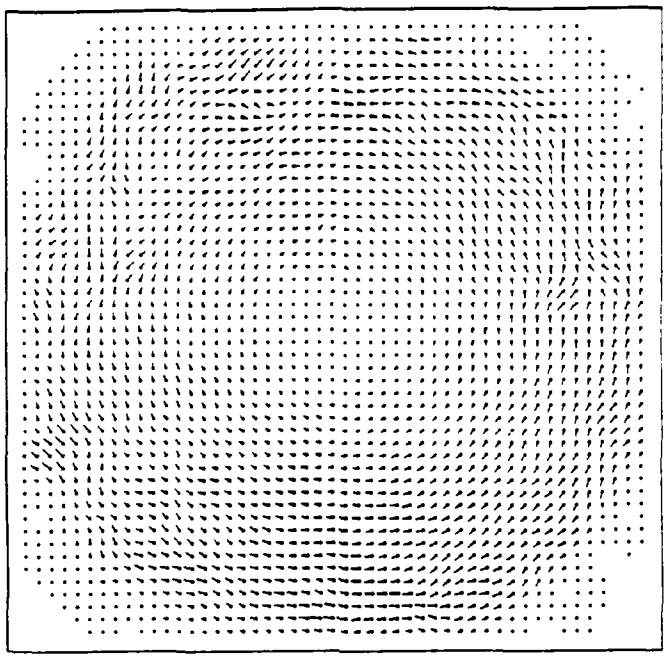

Figure 6: The optical flow field (rotating circle) generated by the local constraint algorithm, followed by Gaussian smoothing (on each component), with $\sigma = 3$.

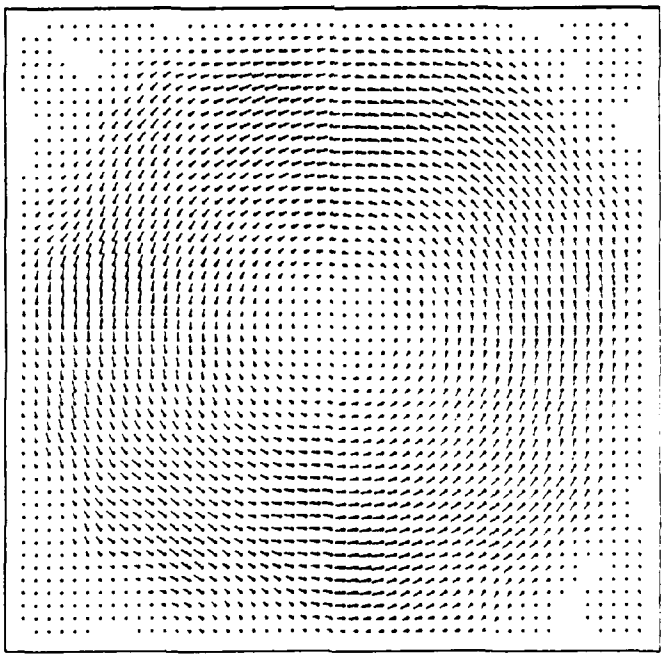

Figure 7: The optical flow field (rotating circle) generated by the gradient constancy algorithm, followed by Gaussian smoothing (on each component), with $\sigma = 3$.

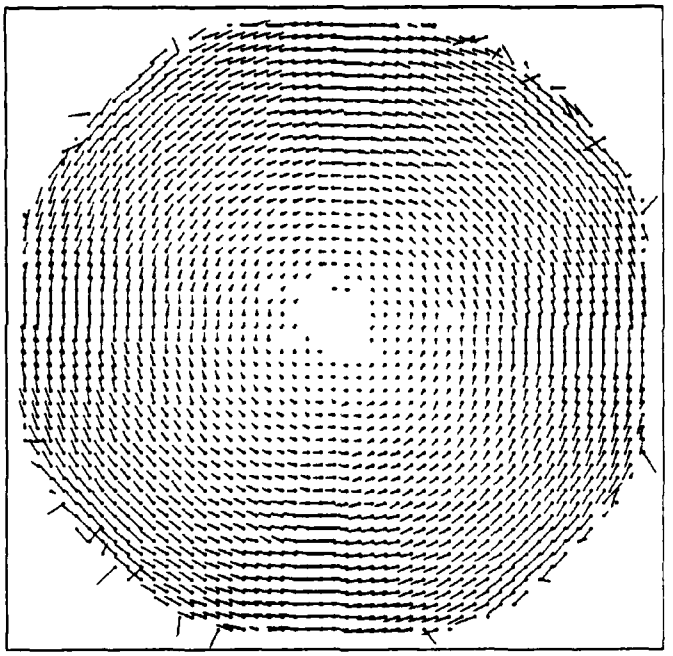

Figure 8: The optical flow field (rotating circle) produced by the matching algorithm using a displacement range of ± 8 pixels.

Algorithm	Aver. $cos\alpha$	Aver. $\|\vec{W}\|$	Aver. of $\frac{\|\vec{W}\|}{\|\vec{V}\|}$
Horn-Schunck (100 steps)	0.976	0.904	0.202
Horn-Schunck (400 steps)	0.977	0.914	0.205
Local constraint	0.992	0.645	0.157
Gradient constancy	0.991	0.744	0.165
Matching (rounded)	0.994	0.196	0.052
Matching (exact)	0.992	0.422	0.129
Matching, interpolated (exact)	0.994	0.252	0.082

Table 1: Comparison of true and computed velocity fields – rotating image.

algorithm generates integer displacements. For the differential algorithms, the exact field $\vec{V}$ is compared with the computed outputs $\vec{V}_c$. The output of the matching algorithm is compared both with the exact field, and the field rounded to integer values. Moreover, the output of the matching algorithm, the matching scores at integral displacements, was interpolated to half pixel increments, resulting in significant improvement in the error measures. Interpolated data for matching reduces the errors both in this example and the next by 35%.

Several measures were computed from this comparison: the average of $cos\alpha$, the cosine of the angle between the true and computed velocity vectors , the average length of $\vec{W} = \vec{V} - \vec{V}_c$, the vector difference between the true and computed velocities and the average of $\frac{|\vec{W}|}{|\vec{V}|}$. Note the slow improvement of the Horn-Schunck algorithm; this flow field and the looming field are both quite smooth, but the regularization stage needs many iterations.

An interesting question is why the performance of differential techniques degrades when moving from translational to rotational velocity fields while matching techniques still produce good results when both are based on similar constraints. The answer is likely to be intimately connected to the nature of differential and matching techniques. Differential techniques like local constraint and gradient constancy need some smoothing of the input data. During the smoothing step, brightness information from locations with different velocities are mixed, introducing error into the differential measurements. The matching algorithm does not, in principle, need smoothing, since it does not take derivatives. Another factor in this and the next example is that a wide range of velocities are present. Differential algorithms depend on the brightness derivatives being well approximated by linear (first derivative) and quadratic terms (second derivative). The images are smoothed by a Gaussian before differentiation, but no one scale of Gaussian suffices to smooth brightness enough at high velocities while not over-blurring locations with small velocities.

Algorithm	Aver. $cos\alpha$	Aver. $\|\vec{W}\|$	Aver. of $\frac{\|\vec{W}\|}{\|\vec{V}\|}$
Horn-Schunck (100 steps)	0.942	0.463	0.321
Horn-Schunck (400 steps)	0.943	0.450	0.314
Local constraint	0.960	0.316	0.174
Gradient constancy	0.977	0.230	0.105
Matching (rounded)	0.988	0.115	0.062
Matching (exact)	0.980	0.405	0.247
Matching, interpolated (exact)	0.992	0.211	0.160

Table 2: Comparison of true and computed velocity fields - looming image.

4.2.2 Looming Field

Let us now consider another simple kind of motion, that of a planar surface translating in space towards the observer (looming). We have compared the performance of the algorithms in the case of a looming planar surface; its appearance is similar to the textured surface used in the preceding example.

The Horn-Schunck algorithm behaves well (see Fig. 9). Differential techniques based on larger support, like local constraint and, in particular, the gradient constancy technique, are not influenced by the spatial structure of the gradient (see Fig. 10 and 11 respectively). The matching algorithm appears not to produce any output in the immediate neighborhood of the singular point of the optical flow — a focus of expansion — because displacements in the neighborhood of the focus of expansion are less than one pixel (see Fig. 12), but does correctly compute zero motion (the rounded approximation to small values) near the singular point.

4.3 Numerical Differentiation

In implementing and testing these algorithms, we directly confronted many of the difficulties in computing derivatives of images. As a concluding point of our analysis, we examine briefly the problems underlying numerical implementation of derivatives of image brightness. Typically, computer vision algorithms — with rare exceptions, for example, [TP86] — use two-point and three-point approximation formulae to compute first and second derivatives of image brightness. For example, an analysis of errors in optical flow gradient methods [KTB87] uses only the forward difference. In essence, it is taken for granted that the inter-pixel distance in space (or average displacements over the image plane between consecutive frames in time) is very "small". A closer analysis, however, shows that this is equivalent to assuming that the filter function used before taking derivatives is oversampled in space (or time), i.e., that it has a "large" scale. To study the effects of various finite difference methods, we use the fact

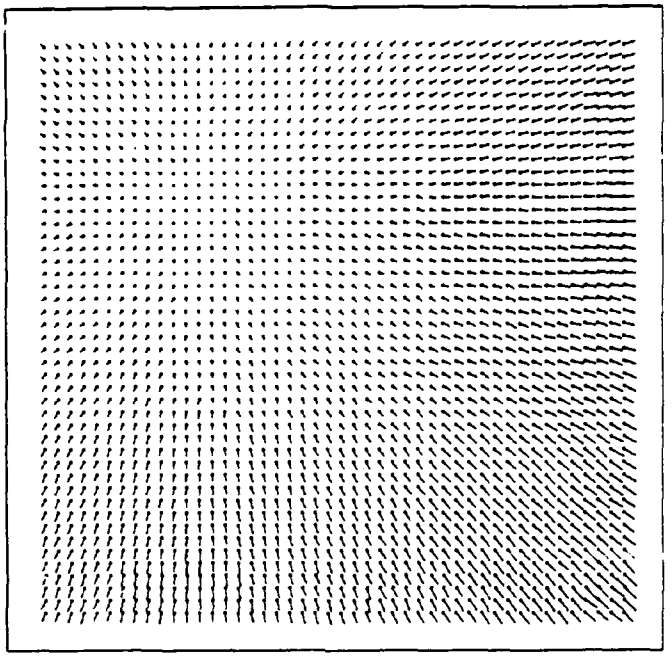

Figure 9: The optical flow (looming image) produced after 400 iterations by the Horn-Schunck algorithm, as in [Hor85], with $\lambda = 1$.

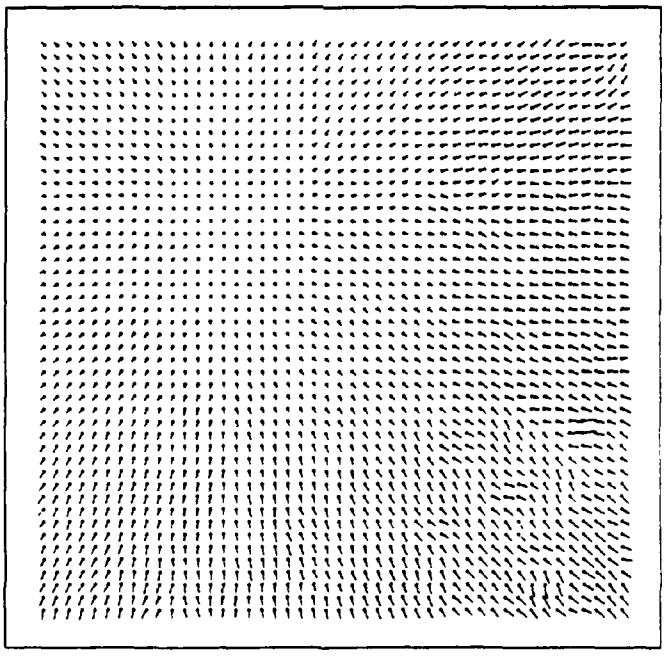

Figure 10: Optical flow (looming image) from the local constraint algorithm. The optical flow is assumed to be constant in an 11 x 11 neighborhood. The vector field is convolved with a symmetric Gaussian function of standard deviation $\sigma = 3$.

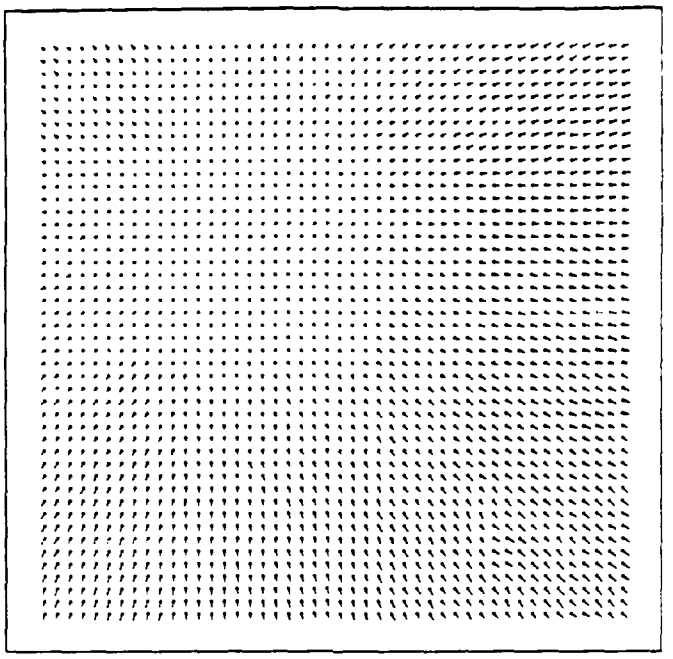

Figure 11: Optical flow (looming image) from the gradient constancy algorithm. The vector field is convolved with a symmetric Gaussian function of standard deviation $\sigma = 3$.

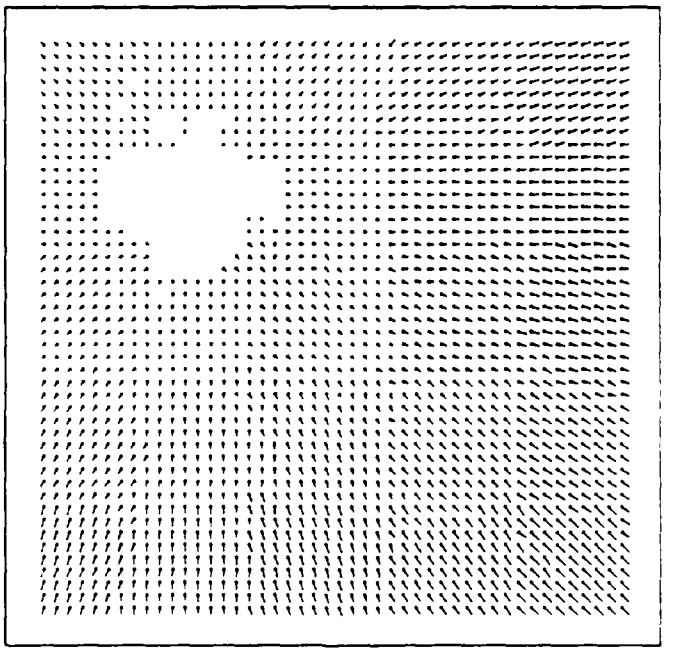

Figure 12: Optical flow (looming image) produced by the matching algorithm using a displacement range of ± 8 pixels.

that differentiation commutes with convolution and examine the approximations to the first and second derivatives of a $1D$ Gaussian generated by convolving the Gaussian with several finite difference operators. Figures 13–16 show some characteristic graphs. The curves plot the true derivative of a $1D$ Gaussian G, and samples of several numerical approximations: the forward or finite difference (two-point formula),

$$f'(x) = f(x+1) - f(x), \tag{18}$$

central or symmetric difference (three-point formula),

$$f'(x) = (f(x+1) - f(x-1))/2, \tag{19}$$

four-point formula,

$$f'(x) = (f(x-1) - f(x+2))/24 + 27(f(x+1) - f(x))/24, \tag{20}$$

and five-point formula,

$$f'(x) = 2(f(x+1) - f(x-1))/3 - (f(x+2) - f(x-2))/12, \tag{21}$$

at the pixel x. In the formulae above, the interpixel distance is unity.

The comparison between the second derivative, f'', of a $1D$ Gaussian function f and numerical approximation of it is obtained by using symmetric difference,

$$f''(x) = f(x+1) - 2f(x) + f(x-1), \tag{22}$$

and five-point support,

$$f''(x) = 4(f(x+1) - 2f(x) + f(x-1))/3 - (f(x+2) - 2f(x) + f(x-2))/12, \tag{23}$$

at the pixel x.

The reference curve in each of the figures is plotted with open circles. For small values of σ, the central difference for the first derivative (squares in Fig. 13) and three-point approximation for the second derivative (triangles in Fig. 15) are clearly insufficient. The error for the approximation A is computed at integer values i for $i < 3\sigma$ and the quantity shown is:

$$\frac{\sum_i^{\lceil 3\sigma \rceil} |A(i) - G'(i)|}{\sum_i^{\lceil 3\sigma \rceil} |G'(i)|} \tag{24}$$

The following table summarizes the errors for various approximations. Note that

σ	2-point	3-point	4-point	5-point	6-point	7-point
1.0	0.0737	0.258	0.0153	0.0875	0.00647	0.0495
1.5	0.0333	0.126	0.00523	0.0339	0.00171	0.0135
2.0	0.0192	0.0745	0.00193	0.0126	0.000380	0.00341
3.0	0.00845	0.0334	0.000392	0.00271	0.0000376	0.000364
4.0	0.00474	0.0188	0.000129	0.000899	0.00000696	0.0000689

Table 3: Normalized summed absolute errors for first derivative of Gaussian.

the error for the central difference only becomes less than 5% for $\sigma >= 3$. This example demonstrates that these estimators are accurate, as theory predicts, when the higher-order derivatives of the function being approximated are small. For the first derivative of the Gaussian, these derivatives become small when $\sigma > 3$ or x becomes large. But, when these conditions are not met, it is clearly superior to use larger support for the approximation. Interestingly, the estimators with an even number of points are consistently better than the odd numbered estimators. For example, at $\sigma = 1$, it is necessary to use the 7-point estimate to achieve accuracy comparable with the 2-point estimate.

But there is a problem with these even number operators. For the first derivative, the estimators with an even number of points, the forward (two-point) and the four-point estimator, are shown displaced leftward by -0.5. Both actually estimate the first derivative at the point $x = 0.5$. For large σ, the relative shift is small. Three- and five-point formulae for the finite difference estimate the derivative at $x = 0$. Any computation which mixes the forward difference operator for the first derivative (which is shifted) with the three-point operator for the second derivative (unshifted) runs into difficulties. The bias of the forward difference operator shows up as a phase shift, dependent on the image frequency, when the frequency response characteristics of the operator are examined.

Let us examine the z-transform of the forward-difference operator $z^{-1} - 1$. The Fourier transform of this operator is

$$z^{-1} - 1\|_{e^{j\omega}} = e^{-j\omega} - 1, \tag{25}$$

which, factored, is

$$e^{-j\omega/2}(e^{-j\omega/2} - e^{j\omega/2}), \tag{26}$$

which reduces to

$$e^{j(\pi/2-\omega/2)}2sin\omega/2. \tag{27}$$

For small angles, $sin\omega$ approximates ω, giving

$$e^{j(\pi/2-\omega/2)}\omega. \tag{28}$$

The Fourier transform of the derivative operator is

$$e^{j\pi/2}\omega. \tag{29}$$

Comparing these two formula we see that the forward-difference operator is an adequate approximation for the derivative operator only for small ω, but, even there, there is a small phase shift, proportional to $\omega/2$. The other operators do not introduce such a phase shift.

These observations explain the symmetric treatment of spatial and temporal derivatives in [Hor85] (p. 289). There the forward difference operator is used, both in space and time. Our analysis indicates that its use must be accompanied by significant smoothing of the image.

Another consequence of the preceding analysis of numerical differentiation is that one can improve the results of implementations of optical flow methods simply by improving the accuracy of the measurement stage. As an example, consider the local constraint method; it does not require subsequent smoothing, so the effects of improved measurements should directly appear in the results. In the results we reported above, we used Horn's symmetric spatial/temporal differentiators. We replaced the x and y components of these operators by the 4-point operators. The results for the rotating image figure show minimal improvement, but other similar images show substantial improvement, reducing the length of the error vector by up to 25%.

Similarly, the gradient constancy utilizes temporal derivatives of dE/dx and dE/dy. The output of that algorithm is most sensitive to the temporal derivatives; changing the operator from the 3-point to the 5-point operator for the first derivative reduces error by 30%.

Numerical differentiation is, in general, ill-posed, and the choice of operator must be influenced by analyis of the task and the noise in the image [TP86]. It is important to note that, while, over space, larger support can be utilized with only a small increase in computation, but over time the situation is rather different, requiring more images. Temporal differentiation is more critical to differential methods than spatial differentiation. The difference in sampling in space and time is significant, and usually several orders of magnitude. In estimation of the temporal derivative, the phase shift induced by the derivative estimators with even numbers of points becomes considerable, given the relatively large spacing between sample points. Considerable care must be taken to handle the estimates appropriately. The symmetric operator used by Horn averages spatial derivates over time and temporal derivatives over space so that estimates are taken at the same time in space-time, resulting in accurate estimates that do not suffer from the phase shift.

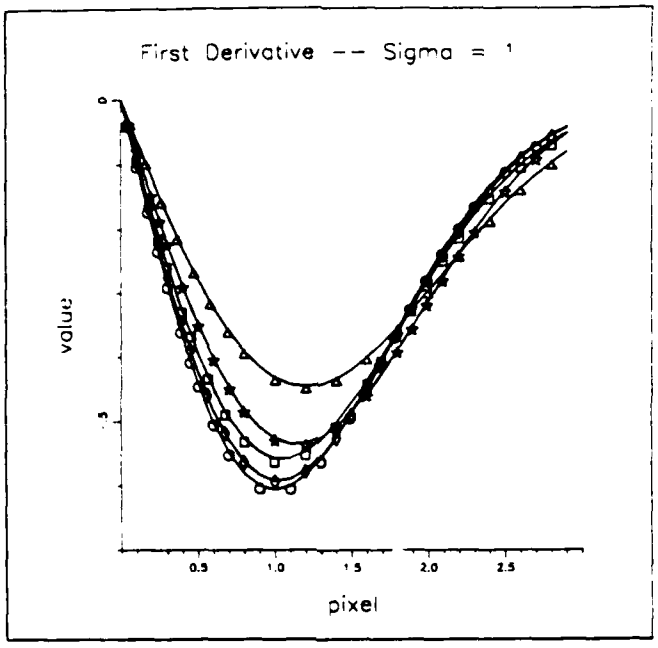

Figure 13: First derivative of Gaussian function ($\sigma = 1$): Gaussian (open circle), forward difference (square), central difference (triangle), four-point (diamond), and five-point (star).

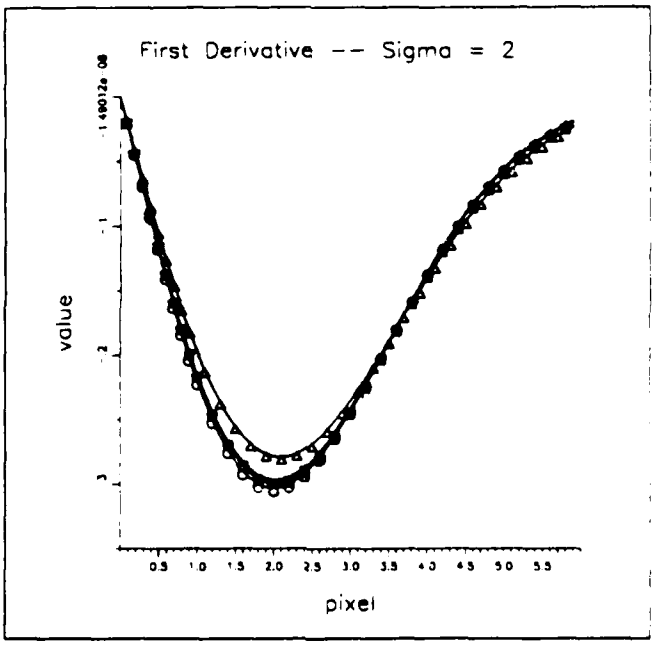

Figure 14: First derivative of Gaussian function ($\sigma = 2$): Gaussian (open circle), forward difference (square), central difference (triangle), four-point (diamond), and five-point (star).

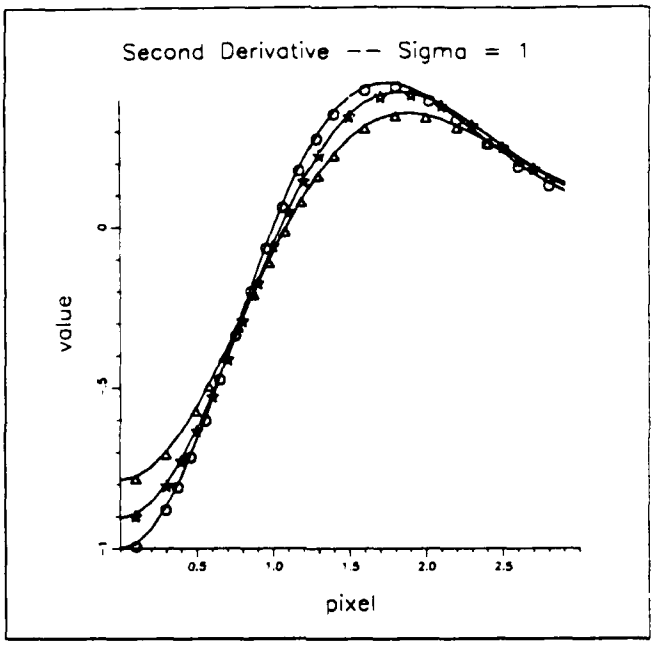

Figure 15: The second derivative of the Gaussian function ($\sigma = 1$): derivative of the Gaussian (open circle); central difference (triangle); five-point formula (star).

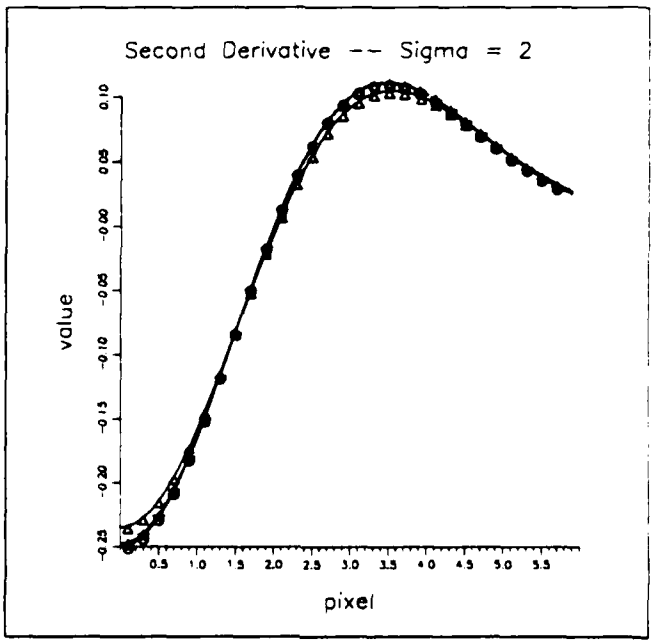

Figure 16: The second derivative of the Gaussian function ($\sigma = 2$): derivative of the Gaussian (open circle); central difference (triangle); five-point formula (star).

5 Conclusion

We have shown that local algorithms can provide good direct measurements of the optical flow, and can locally solve the aperture problem. These algorithms require a simpler, less computationally demanding smoothing stage than global algorithms which begin with normal motion. Since less smoothing is used, localization and discontinuity identification are likely to be better. These improved algorithms require larger spatial support, and thus need to make stronger assumptions about the optical flow. We have shown that, in the cases of pure translation and pure rotation, the projected velocity field which produces the optical flow will meet the local requirements of slow variation except close to singular points of the velocity field.

We have implemented these algorithms on the Connection Machine [LBC89] and tested them on various real and synthetic images, to test their sensitivity to violations of their assumptions. It is clear from both analysis and experience in the implementation that care must be taken in computing derivatives. Interestingly, the methods which produce improved measurements have much in common: they use larger spatial support and they rely on similar assumptions on the local behavior of the velocity field.

References

[Ana87] P. Anandan. A unified perspective on computational techniques for the measurement of visual motion. In *Proceedings Image Understanding Workshop*, pages 219–230, Los Angeles, CA, February 1987. Morgan Kaufmann, San Mateo, CA.

[BH83] A. Bruss and B. K. P. Horn. Passive navigation. *Computer Vision, Graphics, and Image Processing*, 21:3–20, 1983.

[BPT88] M. Bertero, T. Poggio, and V. Torre. Ill-posed problems in early vision. *Proceedings of the IEEE*, 76(8):869–889, August 1988.

[Dev75] P. Dev. Perception of depth surfaces in random-dot stereograms. *Int. Journ. of Man-Machine Studies*, 7(4):511–528, 1975.

[DN82a] L. Dreschler and Hans-Helmut Nagel. On the selection of critical points and local curvature extrema of region boundaries for interframe matching. In *Proceedings Int. Conf. on Pattern Recognition*, pages 542–544, 1982.

[DN82b] L. Dreschler and Hans-Helmut Nagel. Volumetric model and 3-D trajectory of a moving car derived from monocular TV frame sequences

of a street scene. *Computer Vision, Graphics, and Image Processing*, 20:199–228, 1982.

[FT79] C. L. Fennema and W. B. Thompson. Velocity determination in scenes containing several moving objects. *Computer Graphics and Image Processing*, 9(4):301–315, April 1979.

[HL83] Robert M. Haralick and J. S. Lee. The facet approach to optic flow. In L. Baumann, editor, *Proceedings Image Understanding Workshop*, pages 84–93, McLean, VA, 1983. Scientific Applications International Corporation.

[Hor85] Berthold K. P. Horn. *Robot Vision*. MIT Press, Cambridge, Mass., 1985.

[HS81] Berthold K. P. Horn and Brian G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[Kan85] K. Kanatani. Structure from motion without correspondence. In *Proceedings Image Understanding Workshop*, pages 107–116, Miami Beach, FL, December 1985.

[KMJ77] R. E. Kelly, P. R. H. McConnell, and S. J.Mildenberger. The gestalt photomapping system. *Photogrammetric Engineering and Remote Sensing*, 43(11):1407–1417, November 1977.

[KTB87] Joseph K. Kearney, William B. Thompson, and Daniel L. Boley. Optical flow estimation: An error analysis of gradient-based methods with local optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-9(2):229–244, March 1987.

[LB88] James J. Little and Heinrich H. Bülthoff. Parallel computation of optical flow. A.I. Memo No. 929, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 1988.

[LBC89] James J. Little, Guy E. Blelloch, and Todd Cass. Algorithmic techniques for vision on a fine-grained parallel machine. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(3):244–257, March 1989.

[LBP88] James J. Little, Heinrich H. Bülthoff, and Tomaso Poggio. Parallel optical flow using local voting. In *Proceedings of the International Conference on Computer Vision*, pages 454–459, Tarpon Springs, Florida, December 1988. IEEE, Washington, DC.

[LHP80] H. C. Longuet-Higgins and K. Prazdny. The interpretation of a moving retinal image. *Proceedings of the Royal Society of London B*, 208:385–397, 1980.

[LK81] B. D. Lucas and Takeo Kanade. An iterative image registration technique with an application to stereo vision. In *Proceedings IJCAI*, pages 674–679, Vancouver, 1981.

[LOY73] Martin D. Levine, D. A. O'Handley, and G. M. Yagi. Computer determination of depth maps. *Computer Graphics and Image Processing*, 4(4):131–150, October 1973.

[MP76] David Marr and Tomaso Poggio. Cooperative computation of stereo disparity. *Science*, 194(4262):283–287, 15 October 1976.

[MU81] David Marr and Shimon Ullman. Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society of London B*, 211:151–180, 1981.

[Nag83a] Hans-Helmut Nagel. Constraints for the estimation of displacement vector fields from image sequences. In *Proceedings IJCAI*, pages 945–951, Karlsruhe, West Germany, August 1983.

[Nag83b] Hans-Helmut Nagel. Displacement vectors derived from second-order intensity variations in image sequences. *Computer Vision, Graphics, and Image Processing*, 21(1):85–117, January 1983.

[Nag87] Hans-Helmut Nagel. On the estimation of optical flow: Relations between different approaches and some new results. *Artificial Intelligence*, 33:299–324, 1987.

[Nis84] H. K. Nishihara. Practical real-time imaging stereo matcher. *Optical Engineering*, 23(5):536–545, 1984.

[RSE88] Werner Reichardt, R. W. Schlögl, and M. Egelhaaf. Movement detector of the correlation type provide sufficient information for local computation of 2-D velocity field. *Die Naturwissenschaften, in press*, 1988.

[Ter86] Demetri Terzopoulos. Image analysis using multigrid relaxation models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(2):129–139, 1986.

[TH84] R. Y. Tsai and T. S. Huang. Uniqueness and estimation of three dimensional motion parameters of rigid objects with curved surfaces. *IEEE*

Transactions on Pattern Analysis and Machine Intelligence, 6:13–27, 1984.

[TP82] O. Tretiak and L. Pastor. Velocity estimation from image sequences with second order differential operators. In *Proceedings Int. Conf. on Pattern Recognition*, pages 16–19, 1982.

[TP86] Vincent Torre and Tomaso Poggio. On edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(2), 1986.

[UGVT88a] Sergio Uras, Federico Girosi, Alessandro Verri, and Vincent Torre. A computational approach to motion perception. *Biological Cybernetics, in press*, 1988.

[UGVT88b] Sergio Uras, Federico Girosi, Alessandro Verri, and Vincent Torre. A computational approach to motion perception. *Biological Cybernetics*, 60:79–87, 1988.

[Ull83] Shimon Ullman. Computational studies in the interpretation of structure and motion: summary and extension. In J. Beck, B. Hope, and Azriel Rosenfeld, editors, *Human and Machine Vision*. Academic Press, New York, 1983.

[VGT90] Alessandro Verri, Federico Girosi, and Vincent Torre. Differential techniques for optical flow. *Journal of the Optical Society of America A*, 7(5):912–922, 1990.

[Wax84] Allen M. Waxman. An image flow paradigm. In *Proc. Workshop on Computer Vision: Representation and Control*, pages 49–57, Annapolis, MD, 1984. Proceedings of the IEEE.

[WN86] H. Westphal and H. H. Nagel. Toward the derivations of three-dimensional descriptions from image sequences for nonconvex moving objects. *Computer Vision, Graphics, and Image Processing*, 24:302–320, 1986.

[WU85] Allen M. Waxman and Shimon Ullman. Surface structure and 3D motion from image flow: a kinematic analysis. *International Journal of Robotics Research*, 1985.

[YG88] Alan L. Yuille and Norberto M. Grzywacz. A computational theory for the perception of coherent visual motion. *Nature*, 333:71–74, May 1988.