MVA'94 IAPR Workshop on Machine Vision Applications Dec. 13-15, 1994, Kawasaki

DERIVATIVE-BASED OPTICAL FLOW ESTIMATION: CONTROLLED COMPARISON OF FIRST- AND SECOND-ORDER METHODS

Jonathan W. Brandt
School of Information Science
JAIST (Japan Advanced Institute of Science and Technology)
Tatsunokuchi, Nomigun, Ishikawa, Japan 923-12
brandt@jaist.ac.jp

ABSTRACT

A generalization of the first-order, local least-squares optical flow estimation algorithm provides a basis for controlled comparison between the first- and second-order order optical flow estimation methods. The generalization results from the fact that the first-order method performs a computation that is equivalent to that of the second-order method when the input is preprocessed to extract the spatial gradient field and the local neighborhood is reduced to a single point. Using the new generalized framework, the relative performance of the two methods is compared through a series of experiments. The experiments reveal that both methods are systematically biased and that this bias can be reduced by low-pass filtering. The experiments also reveal that the second-order-based optical flow estimates are generally much less accurate than the first-order-based estimates, and that this inaccuracy is due primarily to the fact that the second-order method employs significantly fewer motion constraints per estimate than the first-order method.

1 INTRODUCTION

A large class of optical flow estimation algorithms is based on partial derivative estimates obtained by the use of finite-differencing convolution kernels [1, 4, 7]. This class can be further subdivided into those methods that use first-order partial derivatives and those that use second-order partial derivatives. It has been claimed that the second-order approach is more accurate than the first-order approach because (1) a unique flow can be determined at each point without imposing extra constraints thus obviating the need for neighborhood operations that can blur the instantaneous flow values, and (2) the underlying first-order constancy constraint is not valid in general [7]. However, a recent empirical study found the first-order method originally proposed by Lucas and Kanade [5, 4] to be the best-performing overall [1]. This study compared the results of nine different optical flow estimation techniques, including the second-order method, for a suite of five synthetic and four real image sequences.

The results of this study are not conclusive, however. For instance, one difficulty in evaluating the results of this study is that several critical input image sequence properties such as signal bandwidth, noise power, and flow field properties were not controlled. Also, since each of the candidate algorithms is significantly different from the others, there is no easy way to study the effects of the various algorithm parameters such as conditioning thresholds, and differencing and low-pass kernels.

Thus, it appears to be unresolved whether the first- or the second-order method is preferable. Also, there does not appear to be any clear information concerning the roles of the various parameters that control these algorithms. In an attempt to shed some light on these matters, a generalized algorithm has been developed that unifies the first- and second-order methods. The performance of this algorithm has been tested extensively under a variety of conditions. As a result, several interesting and significant conclusions can be drawn concerning the relative performance of the first- and second-order methods. This paper describes the unified algorithm, the experiment methodology, and presents and evaluates a subset of the experiment results.

2 MOTION CONSTRAINTS

First-order Constraint: Let $E = E(\mathbf{x}, t)$ denote the image intensity function, where $\mathbf{x}$ is a two-dimensional vector specifying a point in the image and t denotes time. Let $\mathbf{v} = [u, v]'$ denote the instantaneous optical flow value. The first-order constancy constraint is

$$\frac{dE}{dt} = \nabla E \cdot \mathbf{v} + E_t = 0, \qquad (1)$$

where $\nabla E = [E_x, E_y]'$ is the intensity gradient relative to the image plane and subscripts denote partial derivatives.

It is well-known that (1) is not sufficient to determine a unique flow value at each point and so additional constraints are required. One of the more successful approaches is to obtain the additional constraints from a finite neighborhood and to combine the constraints by weighted, linear least-squares [5]. Specifically, the *weighted local least-squares method* finds the $\mathbf{v}$ that minimizes

$$\sum_{i=1}^{n} w_i(\nabla E^{(i)} \cdot \mathbf{v} + E_t^{(i)})^2,$$

where $E^{(i)} = E(\mathbf{x}+\Delta\mathbf{x}_i, t+\Delta t_i)$ and w_i is the weight associated with constraint i. The pairs $(\Delta\mathbf{x}_i, \Delta t_i)$ determine a neighborhood around each point from which n first-order constancy constraints are extracted. If the gradient is non-zero and its direction varies sufficiently in the neighborhood, then the resulting system of equations is well-conditioned [4].

If E is vector- rather than scalar-valued, for instance, if it is color, then each vector component contributes a potentially independent set of equations to the linear system. In this case, the neighborhood can be reduced to a single point, while still retaining a system of sufficient rank.

Second-order Constraint: Starting from the first-order constancy constraint, $dE/dt = 0$, and taking the gradient of both sides yields

$$\nabla \frac{dE}{dt} = \frac{d}{dt}\nabla E + M\nabla E = \mathbf{0},$$

where

$$M = \begin{bmatrix} u_x & v_x \\ u_y & v_y \end{bmatrix}$$

is the spatial variation of the flow field. If $\mathbf{v}$ is spatially constant, then the above implies

$$\frac{d}{dt}\nabla E = \mathbf{0}. \qquad (2)$$

Equation (2) is the second-order constancy constraint and is the basis for the optical flow estimation method that has been proposed by Uras et al [6, 7]. From (2) it quickly follows that

$$H_E\mathbf{v} = -\frac{\partial}{\partial t}\nabla E, \qquad (3)$$

where

$$H_E = \begin{bmatrix} E_{xx} & E_{xy} \\ E_{xy} & E_{yy} \end{bmatrix}$$

is the Hessian of E.

If H_E is non-singular, then $\mathbf{v}$ can be determined uniquely without any additional constraints. Because of this, it has been claimed that the second-order constraint is preferable to the under-determined first-order constraint [7]. However, in practice, the accuracy and stability of the second-order method seem to fall short of that of the first-order, local least-squares method [1].

There are several potential reasons why the second-order method is less accurate in practice. One is that it is based on second derivative estimates, which are inherently less accurate than first derivative estimates. A second reason is that only two constraint equations are used per flow estimate, rather than 9 or more. Finally, it is not true in general that $\mathbf{v}$ is spatially constant.

Unification: Suppose $E = \nabla F$, where F is a scalar-valued image intensity function. Since E is vector-valued, it is possible to apply the first-order least-squares estimation method to E using the single-point neighborhood. If the constraints of the two components of E, F_x and F_y, are weighted equally, the least-squares solution is the solution to

$$H_F^2 \mathbf{v} = -H_F \frac{\partial}{\partial t} \nabla F. \qquad (4)$$

The local least-squares method rejects all points for which H_F^2 is ill-conditioned. Hence, it is safe to assume that H_F is non-singular. Consequently, (4) is equivalent to

$$H_F \mathbf{v} = -\frac{\partial}{\partial t} \nabla F.$$

The above implies that the first-order, local least-squares method, when applied to the spatial gradient of an image and using the single-point neighborhood, performs a computation that is equivalent to that of the second-order method. It also suggests that it may be possible to blend the best features of the first- and second-order methods. In order to achieve such a blending, it is necessary to understand more thoroughly the strengths and weaknesses of the two methods. Therefore, some experimentation is required. The new, unified algorithm is an excellent framework for experimentation because it enables controlled comparison between the two methods. In the next section, some preliminary experimental results, based on this approach, are presented.

3 EXPERIMENTS

The experiment methodology is as follows. A synthetic two-dimensional image is generated by isotropically filtering a white Gaussian noise field so that the resulting image has a flat spectrum from zero up to 20% of the Nyquist frequency and zero energy above this cut-off. The image is then rotated about its center by two degrees per frame to generate an image sequence. The resulting sequence (179 frames of 128×128 pixels each) is then presented as input to the generalized optical flow estimation algorithm. The mean and standard deviation of the estimated velocity magnitude and direction at each pixel are gathered by integrating the results of processing the entire sequence.

The local neighborhood for the first-order method is the 3×3 neighborhood generated by the kernel $[.25, .5, .25]$, while for the second-order method it is the single-point neighborhood. In both cases, the differencing kernel is $[-1, 8, 0, -8, 1]/12$ and the conditioning threshold for the linear system is 20. (Flow estimates for which the conditioning number is greater than 20 are rejected.)

The performance of the first- and second-order methods can be compared by examining Figure 1. The plots depict the flow magnitude estimation bias and standard deviation for the first- and second-order methods as functions of the actual flow magnitude. Under these conditions, it appears that the first-order method is more accurate than the second-order method: it produces estimates with slightly less bias and significantly less variance than those of the second-order method. In both cases, the bias is generally positive for velocity magnitudes less than one and negative for velocity magnitudes greater than one. This is probably due to the non-ideal response of the differencing kernel — a phenomenon that has been described elsewhere [3, 2]. The very high bias in the second-order case at low velocities may be due to the spatial variation of the velocity field near the center of rotation. Repeating the experiment with a reduced angular increment per frame reduces this bias, and consequently supports this hypothesis.

Nevertheless, the second-order method does not appear to be very useful because of the high variance in the resulting estimates. The question is, what is the source of this high variance and how can it be reduced? It is unlikely that it is caused by spatial variation in the velocity field because the variance is larger for points that are farther from the center of rotation.

It is often stated that the input to derivative-based optical flow estimation algorithms must be "regularized" by low-pass

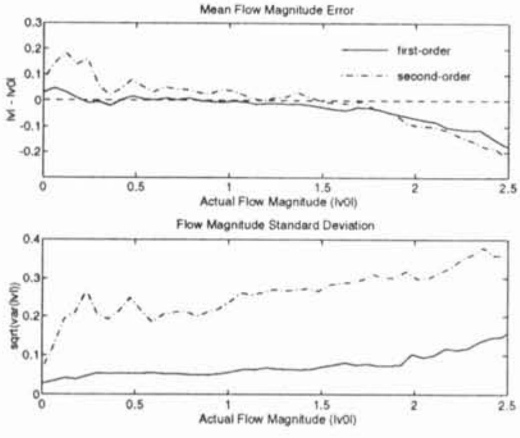

Figure 1: The mean error and standard deviation of the flow velocity magnitude estimates are plotted as functions of the actual flow velocity magnitude. The input is filtered white Gaussian noise (see text). All quantities are in units of pixels/frame.

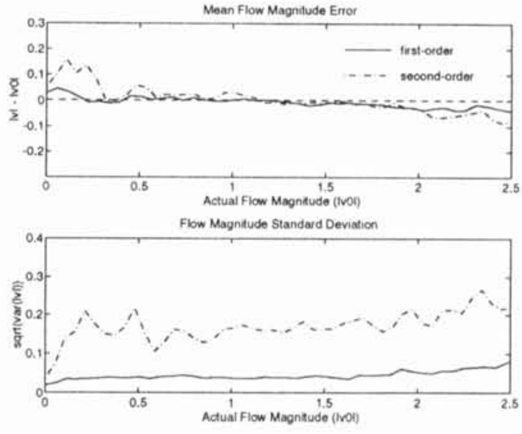

Figure 2: The first experiment is repeated, but with the input first low-pass filtered using a $9 \times 9 \times 9$ point Gaussian ($\sigma = 1.5$).

filtering due to the ill-posedness of numerical differentiation. So, to determine the effect of regularization on these estimates, the first experiment is repeated, but with the input sequence first low-pass filtered using an isotropic $9 \times 9 \times 9$ point Gaussian kernel ($\sigma = 1.5$). The results are shown in Figure 2. It is clear that the performance of both methods has been improved by pre-filtering. In particular, the systematic bias has been greatly reduced.

Additional experiments reveal that it is the broad temporal support of the low-pass kernel that contributes most significantly to the improvement in accuracy. Since the flow field for these experiments is constant with respect to time, there is no penalty in using a low-pass kernel with very broad temporal support. In real applications, the temporal support of the low-pass kernel is limited by the required temporal resolution of the optical flow estimates.

Despite the dramatic accuracy improvement resulting from low-pass filtering, the variance resulting from the second-order method is still unacceptably high and thus apparently not directly due to a lack of regularization, but instead to other sources. For instance, it is possible that the conditioning threshold is set too high and that this is the source of the high variance in the estimates resulting from the second-order method. However, repeating the first experiment with the conditioning threshold reduced to 10, and then to 6, produces no significant change in the results. The conditioning threshold cannot be set any lower because the estimates become too sparse to collect meaningful statistics. Thus, it appears that the conditioning threshold is not the source of the high variance.

One significant difference between the two methods, which was alluded to in Section 2, is that the first-order method uses nine constraint equations (with the 3×3 neighborhood) while the second-order method uses only two. There is no intrinsic reason why the second-order method cannot use a larger neighborhood to overconstrain the flow and thereby stabilize the estimates. To test the effect of overconstraining the second-order-based estimates, the first experiment is repeated, but with the local neighborhood for least-squares estimation for the second-order case set to the same 3×3 neighborhood as that for the first-order case. The results are shown in Figure 3.

It is evident from the data in Figure 3

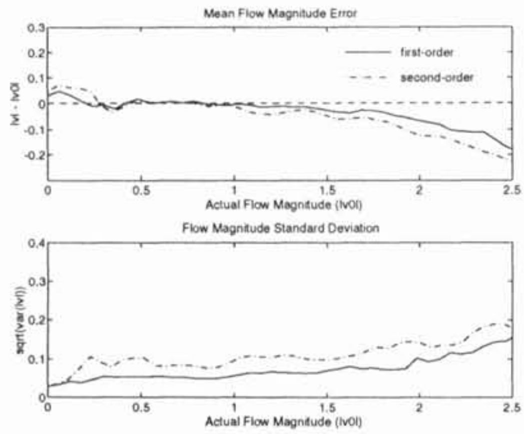

Figure 3: The first experiment is repeated, but in this case the local neighborhood for least-squares estimation for the second-order case is the same as that for the first-order case (3×3).

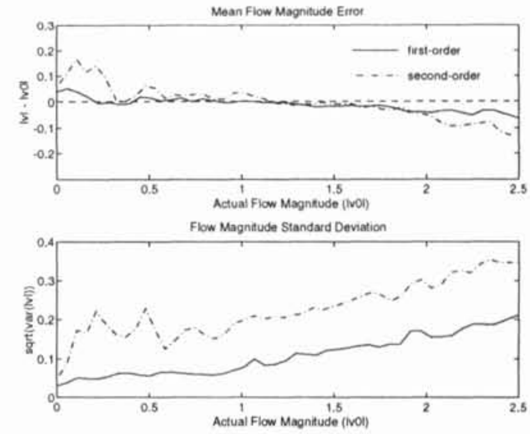

Figure 4: The second experiment is repeated, but with white noise added to the input (SNR = 13).

that overconstraining the second-order-based estimates significantly improves their accuracy. The bias has been reduced slightly and the variance has been reduced significantly. In fact, the performance of the second-order method is close to, but not quite as good as, that of the first-order method in this case. These data tend to undermine the supposed advantage of the second-order method of being able to evaluate the optical flow at a point, without additional information, since the resulting estimates are highly inaccurate, mostly because the estimates are based on minimal information.

In the experiments described above, the input is essentially noise-free, except for a small amount due to interpolation. The noise-free condition is very useful in identifying systematic error sources in the two methods. Nevertheless, it is vital to compare the relative performance of the two methods in the presence noise. When the first experiment is repeated with white noise added to the input, the estimates resulting from both methods are extremely inaccurate, both in terms of bias and variance. However, if the input is low-pass filtered, as in the second experiment, then both methods appear to be relatively noise-tolerant.

Figure 4 depicts the case where the input is corrupted by additive white noise (SNR = 13), and then low-pass filtered prior to optical flow estimation. Under these conditions, the estimation bias resulting from the two methods is unaffected, but the variance is increased, as one would expect. It is interesting to note that the change in variance with additive white noise increases with velocity magnitude. That is, higher velocity estimates are relatively less reliable in the presence of noise than lower velocity estimates. Finally, neither method appears to be more or less sensitive to noise than the other, provided that low-pass filtering is performed as a preprocessing step.

4 CONCLUSIONS

In summary, a new unification of first- and second-order derivative-based optical flow estimation provides a framework for controlled comparison of the two methods. The experiments reported here reveal that both methods are systematically biased and that this bias can be reduced by low-pass filtering. The experiments also reveal that the second-order-based optical flow estimates are generally much less accurate than the first-order-based estimates, and that this inaccuracy is due primarily to the fact that the second-order method employs only two motion constraints per estimate while the first-order method generally employs nine or more constraints per estimate. When the number of constraints employed by the second-order method is increased, by increasing the local, least-squares neighborhood, its accuracy improves and approaches that of the first-order method. However, in none of the experiments described here has the second-order method been as accurate as the first-order method.

More experiments are needed to completely determine the conditions under which the first- or the second-order method is preferable. In particular, it is necessary to examine the roles of the other algorithm parameters. Also, it is very important to examine more closely the claim that the second-order method is immune to violations in the first-order constancy constraint. Within the experimental framework described here, it appears to be a simple matter to generate more realistic sequences where this assumption is violated. For instance, oblique views with time-varying illumination can be simulated so as to test this claim. In any event, the unified first- and second-order optical flow algorithm is an excellent basis for future experimentation and application.

ACKNOWLEDGEMENT

This research was undertaken at JAIST under a grant from Komatsu, Ltd.

References

[1] J. L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *Int. J. Computer Vision*, 12(1):43–77, 1994.

[2] J. W. Brandt. Analysis of bias in gradient-based optical-flow estimation. In *Proc. Twenty-Eighth Annual Asilomar Conference on Signals, Systems, and Computers*, 1994. In press.

[3] J. W. Brandt. Finite-differencing errors in gradient-based optical flow estimation. In *Proc. First IEEE Int. Conf. Image Processing*, 1994. In press.

[4] J. K. Kearney, W. B. Thompson, and D. L. Boley. Optical flow estimation: an error analysis of gradient-based methods with local optimization. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 9(2):229–244, March 1987.

[5] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proc. 5th Int. Joint Conf. Artificial Intelligence*, pages 674–679, 1981.

[6] S. Uras, F. Girosi, A. Verri, and V. Torre. A computational approach to motion perception. *Biological Cybernetics*, 60:79–87, 1988.

[7] A. Verri, F. Girosi, and V. Torre. Differential techniques for optical flow. *J. Optical Soc. America A*, 7(5):912–922, May 1990.