Analysis of Bias in Gradient-Based Optical-Flow Estimation

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

Accurate gradient-based optical-flow estimation depends on accurate partial derivatives that are generally approximated by the use of finite-differencing convolution kernels. The consequent error in the first derivative estimate is approximately proportional to the third derivative of the input signal and leads to systematic errors in the optical-flow estimates. Simulations indicate that these errors tend to dominate other error sources, such as broad-band noise, unless the system is carefully tuned. The result suggests that the high-frequency attenuation of the finite-differencing kernel imposes a resolution limit on the estimated optical-flow field.

1 Introduction

Accurate estimation of the optical-flow field of an image sequence is a long-standing and important problem in computer vision. Accuracy is especially important when the optical-flow field is provided as input to numerically sensitive three-dimensional motion and structure estimation algorithms. In this case, systematic estimation errors can lead to disaster by biasing critical motion parameters such as time-to-impact.

A large class of optical-flow estimation algorithms rely on accurate partial derivative estimates obtained by the use of finite-differencing convolution kernels [1, 5, 7]. Such kernels can accurately approximate the $j\omega$ frequency-response characteristic of the derivative operator only over a limited frequency range. This frequency range is related to the kernel's support size and thus ultimately to the operating speed and cost of the system that performs the computations.

It is usually recognized that spatio-temporal low-pass filtering of the input image sequence is necessary in order to obtain accurate results. The need for low-pass filtering has generally been attributed to the presence of broad-band noise in the input. In this paper, it is proposed that noise is, in many cases, not the predominant contributor to optical-flow estimation errors. Instead, the non-ideal response of the differentiator presents a fundamental obstacle to accurate gradient-based optical-flow estimation.

2 Analysis of bias

Kearney's error analysis [5] identified the gradient estimation step as a source of systematic error for optical-flow estimation. Using the Taylor-series expansion, he argued that the derivative of a function $f(x)$ that is estimated by

$$\hat{f}'(x) = \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

yields the approximation

$$\hat{f}'(x) \approx f'(x) + \Delta x f''(x).$$

So the error in the derivative estimate is proportional to the second derivative. However, it is generally better (and more common) to use central differences to estimate the derivative. That is,

$$\hat{f}'(x) = \frac{f(x+\Delta x) - f(x-\Delta x)}{2\Delta x},$$

which effectively cancels the second-order term in the Taylor-series expansion and yields the alternative approximation

$$\hat{f}'(x) \approx f'(x) + \frac{\Delta x^2}{6} f'''(x).$$

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

1058-6393/95 $4.00

In practice, central differencing produces errors that are well-characterized by the latter formula — the derivatives are generally underestimated when the curvature is decreasing, overestimated when it is increasing, and not significantly affected by the value of the second derivative. Let the positive parameter α specify the efficacy of the differentiator in estimating the derivative of f,

$$\hat{f}'(x) \approx f'(x) + \alpha f'''(x). \qquad (1)$$

The above approximation is also collaborated by frequency-domain analysis. Suppose the frequency-response characteristic of a finite differencing kernel is modeled as

$$D(\omega) = j\omega A(\omega), \qquad (2)$$

where $A(\omega)$ is the frequency-response characteristic of a low-pass filter. Then, in the frequency domain, the difference between the ideal and the estimated derivative is

$$\hat{F}' - F' = F'(A-1) = F'''(1-A)/\omega^2.$$

So, the approximation in (1) is equivalent to assuming

$$A(\omega) \approx 1 - \alpha\omega^2, \qquad (3)$$

which is a very reasonable form for a low-pass filter. For instance, the Gaussian low-pass filter has a frequency-response characteristic that is proportional to $e^{-\alpha\omega^2}$, which has the Taylor-series expansion

$$e^{-\alpha\omega^2} = 1 - \alpha\omega^2 + O(\omega^4).$$

Therefore, the differencing kernel derived from the first derivative of the Gaussian is subject to errors of the form expressed in (1). Many other differencing kernels are subject to these errors as well.

Now, consider the one-dimensional optical-flow estimation problem: given a signal of the form

$$f(x,t) = g(x - vt)$$

that has been sampled in space and time, estimate v. It is usually assumed that v varies slowly relative to x and t. By applying the one-dimensional flow-constraint equation,

$$f_x v + f_t = 0,$$

(subscripts denote partial derivatives) the flow velocity can be estimated (provided $\hat{f}_x \neq 0$) by

$$\hat{v} = -\hat{f}_t/\hat{f}_x.$$

Substituting the approximation formulae for $\hat{f}_t$ and $\hat{f}_x$ yields

$$\hat{v} \approx \left(\frac{g' + \alpha v^2 g'''}{g' + \alpha g'''}\right) v.$$

Thus, the relative bias in the flow velocity estimation is

$$(\hat{v} - v)/v \approx \left(\frac{\alpha g'''}{g' + \alpha g'''}\right)(v^2 - 1). \qquad (4)$$

The above formula indicates that the relative bias is directly proportional to α and g''', and inversely proportional to g'. The term $(v^2 - 1)$ is intriguing because it implies that the bias goes to zero as $|v|$ approaches unity.

In the frequency domain, it is not difficult to show that

$$F(\omega_1, \omega_2) = G(\omega_1)\delta(\omega_2 + v\omega_1)$$

which implies that the temporal frequencies are scaled by $|v|$ relative to the spatial frequencies. (In fact, this relationship is the basis for Heeger's frequency-domain approach to optical-flow estimation [4].) If the spatial and temporal derivatives are each approximated by the same differencing operations $D(\omega)$, then the consequent distortion in the temporal domain will generally differ from that in the spatial domain. However, if $|v|$ is close to unity, then the distortions in the temporal and the spatial domains will be nearly identical, and so the errors in $\hat{f}_t$ and $\hat{f}_x$ will more-or-less cancel. This explains why the error approaches zero as the flow velocity approaches unity, regardless of how poorly the system operates (that is, regardless of the value of α). Thus, when evaluating such an estimation system, it is very important to test its accuracy over a broad velocity range.

It has been shown elsewhere [3] that if $g(x) = sin\omega x$ and the differentiator is of the form expressed in (2), then the relative flow estimation bias is

$$(\hat{v} - v)/v \approx \left(\frac{A(v\omega)}{A(\omega)} - 1\right).$$

If $A(\omega)$ is of the form specified in (3), then the above expression is equivalent to

$$(\hat{v} - v)/v \approx \left(\frac{\alpha\omega^2}{1 - \alpha\omega^2}\right)(1 - v^2), \qquad (5)$$

which can be seen as the frequency-domain alternative to (4) that applies when $g(x)$ is a sinusoid, or more generally, a narrow-band signal. The term $\alpha\omega^2/(1-\alpha\omega^2)$ is non-negative, so $\text{sgn}(\hat{v} - v) = \text{sgn}((1 - v^2)v)$. This is significant because it suggests that for a narrow-band signal, gradient-based optical-flow magnitude estimates are systematically biased toward unity.

3 Simulations

A series of simulations was conducted in order to test the hypothesis that the non-ideal response of the differentiator systematically biases optical-flow estimation. The procedure for each simulation run is as follows. First, a two-dimensional image with prescribed spatial frequency content is generated. The image is an instance of independent, identically-distributed white Gaussian noise that has been passed through a filter with frequency-response characteristic

$$H_{\rho_1,\rho_2}(\omega_1,\omega_2) = \begin{cases} 1 & \text{if } \rho_1 \leq \sqrt{\omega_1^2+\omega_2^2} \leq \rho_2 \\ 0 & \text{otherwise.} \end{cases}$$

This filter is an isotropic band-pass filter whose pass band is delimited by the parameters ρ_1 and ρ_2. Three images are generated using the parameters $(\rho_1, \rho_2) = (0.0, 0.2)$, $(0.2, 0.4)$, and $(0.4, 0.6)$, respectively. (These frequencies are normalized such that the Nyquist frequency is one.) These parameters were chosen so that it would be possible to compare the relative optical-flow estimation error resulting from information originating in different spatial frequency regions. Let these three images be called the ***low-***, ***medium-***, and ***high-frequency*** images, respectively.

Each of these three images is then rotated about its center by a fixed angular increment per frame to generate an image sequence that specifies a rotating motion field. This rotating motion field has the property that the velocity at each pixel is unique and constant over time. In the experiments reported here, the image is rotated by 2° per frame and the resulting sequence is 179 frames of 128 × 128 pixels each.

Each synthetic image sequence is then processed to estimate the optical-flow by the method of local least-squares [6, 5, 1]. The resulting optical-flow estimates at each pixel are integrated over time to collect mean and variance statistics. Figure 1 depicts the relative mean error and relative standard deviation in the flow velocity magnitude estimates as functions of the actual magnitude. In this case, the differencing kernel was $[-1, 8, 0, -8, 1]/12$, which is the kernel suggested in [1].

It is clear from Figure 1 that the optical-flow is more accurately estimated from the low-frequency image than from the medium- and high-frequency images. It should be noted that the aliasing limit for flow magnitude estimation is 1.67 pixels per frame for the high-frequency image and 2.5 pixels per frame for the low-frequency image. Nevertheless, significant estimation bias occurs for these images well within the non-aliasing region. It is also clear that the flow

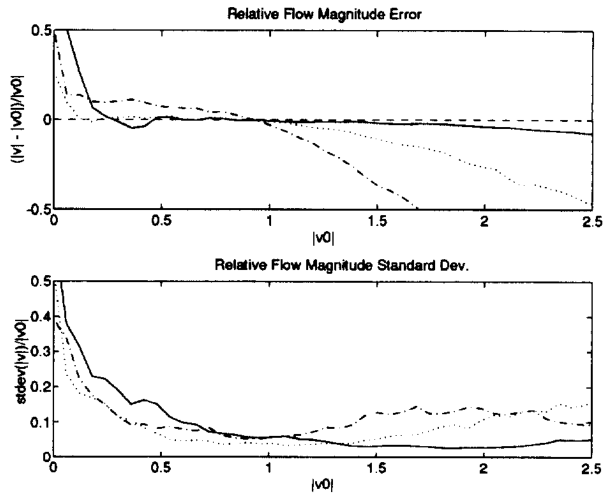

Figure 1: The relative mean error and relative standard deviation of the flow velocity magnitude estimates are plotted as functions of the actual flow velocity magnitude. The inputs are filtered white Gaussian noise (see text). The filter parameters are $(0.0, 0.2)$ (solid curve), $(0.2, 0.4)$ (dotted curve), and $(0.4, 0.6)$ (dash-dotted curve). All quantities are in units of pixels per frame.

magnitude is systematically under-estimated when it is greater than one, as predicted under the narrow-band assumption in Section 2. However, the predicted over-estimation of flow magnitudes less than one is not clearly observed, except in the high-frequency case. This may be because the narrow-band assumption applies only to the high-frequency image.

The relative standard deviation in the flow magnitude estimates is somewhat high considering that there is essentially no noise in the input, except for a small amount due to interpolation of the rotated images and due to slight local variation in the optical-flow. The relative standard deviation is is especially high at lower velocities. The source of the variance in the estimates is most likely the finite bandwidth of the input that distorts the estimates as predicted in (5). Independent experiments with narrow-band one-dimensional signals resulted in essentially zero variance in optical-flow estimation, even when the mean estimates were highly biased [3].

The preceding experiment demonstrates that the optical-flow estimation bias actually occurs in practice. The question is, how dominant is this error? Does it dominate other sources, such as broad-band noise in the input? To investigate this question, the same experiment was repeated, but with a moderate level of white noise added to corrupt the input (SNR = 18dB). The results of this experiment are depicted in Figure 2. At this noise level, the flow magnitude estimates resulting from the low-frequency image have deteriorated relative to those of the noise-free case. However, the estimates derived from the medium- and high-frequency images are relatively unaffected. This implies that at the medium- and high-frequencies, the systematic distortion is still predominant. When the noise is increased to the level such that SNR = 10dB, the flow estimates resulting from the medium- and high-frequency images begin to deteriorate as well.

Since the previous experiments indicate that medium- and high-frequency information in the input tends to corrupt the flow estimates, it appears necessary to low-pass filter the input as a preprocessing step. In order to test this hypothesis, the first experiment was repeated, but with the input pre-filtered with a $9 \times 9 \times 9$ Gaussian kernel ($\sigma = 1.5$). The results are shown in Figure 3.

It appears that the estimates are much more reliable when the input is low-pass filtered. The most marked change is in the medium-frequency case. The spatial frequency range of this image is in the transition region of the low-pass filter that was used in the

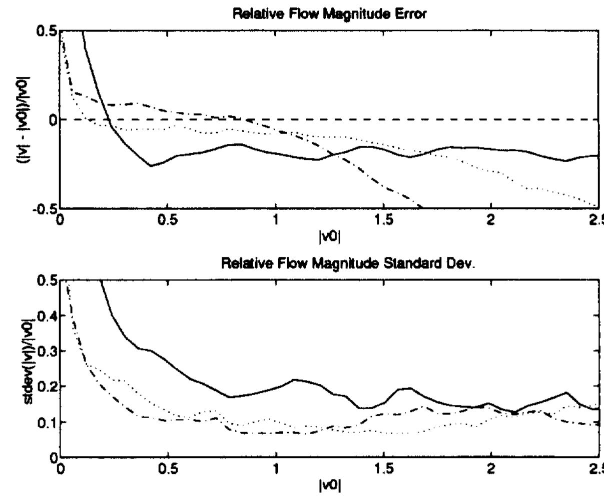

Figure 2: The relative mean error and relative standard deviation of the flow velocity magnitude estimates when moderate white noise is added (SNR = 18dB). The inputs are identical to those in Figure 1.

experiment. It is especially significant that the variance of medium-frequency-based estimates has been dramatically reduced. Although the mean estimates for the high-frequency case have also been improved, the variance in this case is quite high. In fact, it has increased significantly over that resulting from the first experiment. This makes sense because most of the high frequency information has been removed by the low-pass filter.

In summary, these simulations appear to confirm the existence of systematic flow-estimation errors of the form predicted in (4) and (5). The errors appear to be particularly severe when the velocity magnitude is greater than one and in general, for high-frequency information.

4 Discussion

One obvious way to minimize this systematic flow-estimation error is to stipulate a maximum velocity magnitude V_0 that can be estimated reliably and to use a low-pass filter that has a cutoff at Ω in the spatial frequency domain and ΩV_0 in the temporal frequency domain. In this case, it is also necessary that the differentiator have a nearly ideal characteristic up to the frequency $\max(\Omega, \Omega V_0)$. Velocity magnitudes greater than V_0 are likely to have a large residual re-

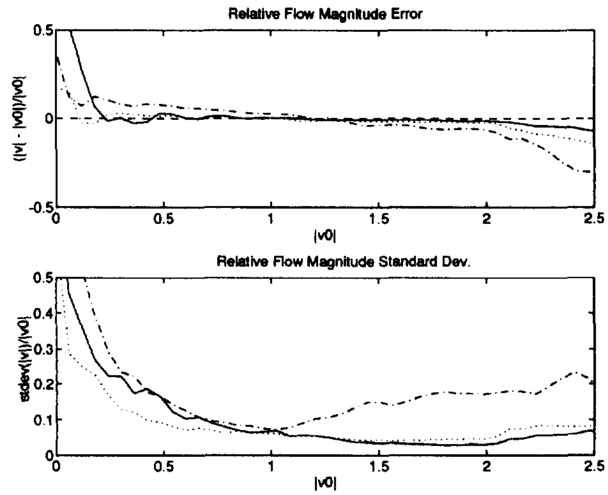

Figure 3: The relative mean error and relative standard deviation of the flow velocity magnitude estimates when the input is low-pass filtered prior to processing. The low-pass filter is the $9 \times 9 \times 9$ point Gaussian kernel with $\sigma = 1.5$. The inputs are identical to those in Figure 1.

sulting from the least-squares fit and therefore can be rejected reliably.

An alternative approach to controlling this error is to use third derivative estimates to indicate the reliability of the first derivative estimates. This approach would fit well in the context of the weighted local least-squares approach to optical-flow estimation [5], where the flow-constraint equations in a local neighborhood are combined to form the estimate. In this case, the third derivative estimates could be used to determine the weights for the least-squares solution — a smaller third derivative magnitude would lead to a greater weight for the corresponding flow-constraint equation. The drawback of this approach is that the third derivatives estimates themselves are subject to similar (or more severe) errors and are thus unreliable. Moreover, additional computation is required. The viability of this approach is under investigation.

5 Conclusions

In this paper, the effect of non-ideal differentiation on gradient-based optical-flow estimation has been explored through analysis and experimentation. The results indicate that the non-ideal differentiator response imposes a limit on the amount of high-frequency information that is useful for optical-flow estimation. Specifically, it has been found that high-frequency attenuation of the differentiator leads to severe distortions in the estimates resulting from high-frequency information. Consequently, high-frequency information must be eliminated by low-pass filtering in order to overcome this problem. This implies that the spatial resolution of the optical-flow field is ultimately limited by the differentiator's frequency-response characteristic.

It is expected that a similar limitation exists for the second-order method [7]. Some preliminary experimentals that compare the systematic errors of the first- and second-order optical flow estimation methods are reported in [2].

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. Derivative-based optical flow estimation: Controlled comparison of first- and second-order methods. In *IAPR Workshop on Machine Vision Applications*, 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] D. J. Heeger. Model for the extraction of image flow. *J. Optical Soc. America A*, 4(8):1455–1471, August 1987.

[5] 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.

[6] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proc. DARPA Image Understanding Workshop*, pages 121–130, 1981.

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