Short Papers

The Performance of Camera Translation Direction Estimators From Optical Flow: Analysis, Comparison, and Theoretical Limits

A. Mark Earnshaw and Steven D. Blostein

Abstract—A noniterative method using optical flow to recover the translation direction of a moving camera has been previously proposed in [4]. We present a detailed explanation of the bias in this algorithm and compare methods for eliminating this bias, as well as presenting a comprehensive error analysis. This analysis includes a necessary modification to the Cramér-Rao lower bound (CRLB). We propose a simple iterative modification to the algorithm which produces unbiased translation direction estimates that approach the CRLB. Numerical results are used to compare the various techniques on synthetic and real image sequences.

Index Terms—Translation direction estimation, linear constraints, optical flow, error analysis, performance comparison.

✦

1 Introduction

One approach for recovering the motion parameters of a mobile camera in a stationary environment uses the observed optical flow data. Heeger and Jepson [4] illustrated how the estimation of translation, rotation, and depth could be decoupled into separate problems which could be solved in order. They later reformulated the original iterative translation recovery method as a noniterative problem [6]. However, the resulting translation direction estimates tended to be biased towards the optical axis. Jepson and Heeger [7] performed a preliminary bias analysis and proposed a solution for eliminating it.

This paper presents a detailed error analysis of the bias which includes alternative methods for overcoming it. We explain why bias compensation is not as straightforward as initially thought since asymptotically unbiased estimators are not necessarily unbiased in the case of a finite number of samples. The error analysis contained here is of significance since it is applicable to general least-squares problems in other research areas.

There are other methods for recovering camera motion parameters such as the investigation of an optimal motion and structure estimation algorithm in [13]. However, their technique is feature-based rather than using optical flow. In addition, their optimal method requires the iterative optimization of a motion vector which is computationally expensive. Conversely, the techniques compared in this paper are noniterative in nature, with the exception of our "optimized" algorithm which only requires the solution of a single variable.

• *The authors are with the Department of Electrical and Computer Engineering, Queen's University, Kingston, Ontario, Canada K7L 3N6.*
E-mail: {earnshaw, sdb}@ee.queensu.ca.

Manuscript received Feb. 3, 1995; revised Feb. 12, 1996. Recommended for acceptance by G. Medioni.
For information on obtaining reprints of this article, please send e-mail to: transpami@computer.org, and reference IEEECS Log Number P96020.

2 Motion Recovery Algorithms

2.1 Optical Flow Observations

The instantaneous optical flow at a point on the image plane of a moving camera is:

$$u_i = \frac{1}{Z_i} \mathrm{A}_i \mathrm{T} + \mathrm{B}_i \Omega$$

$$\mathrm{A}_i = \begin{bmatrix} -f & 0 & x_i \\ 0 & -f & y_i \end{bmatrix} \quad \mathrm{B}_i = \begin{bmatrix} x_i y_i & -\left(1+x_i^2\right) & y_i \\ \left(1+y_i^2\right) & -x_i y_i & -x_i \end{bmatrix} \tag{1}$$

where T and Ω represent the translational and rotational velocities of the camera, (x_i, y_i) specify the image plane coordinates of the ith pixel, and Z_i is the corresponding unknown depth. Note that A_i and B_i are completely defined by the camera geometry.

Our measurement noise model assumes additive Gaussian noise with an identical distribution over the entire image. These terms are assumed to be independent with identical Gaussian distributions $N(0, \sigma_n^2)$. The standard deviation of the noise σ_n is a specified percentage ρ of the average flow vector length and is assumed constant over the image.

2.2 Original Biased Method (Jepson and Heeger)

The observed optical flow may be used to derive linear constraint vectors τ_i which are (in the noiseless case) perpendicular to the translation direction T. These τ_i vectors are calculated as a convolution over a local patch of twisted flow vectors, $\mathbf{q}_j$. Each $\mathbf{q}_j$ is the cross-product of an optical flow vector $\mathbf{u}_j$ with the corresponding sampling direction vector $\mathbf{s}_j$.

$$\tau_i = \sum_j c_{ij} \mathbf{q}_j \tag{2}$$

$$\mathbf{q}_j = \mathbf{u}_j \times \mathbf{s}_j \tag{3}$$

$$\mathbf{s}_j = \begin{bmatrix} x_j & y_j & f \end{bmatrix} \tag{4}$$

(x_j, y_j) specify image plane coordinates, and f is the focal length of the camera lens. The c_{ij} represent a set of 2D filter coefficients, centered at position i, whose sum-of-squares is normalized to unity. These coefficients are selected to ensure that the constraints remain perpendicular to the translation direction in the noiseless case regardless of the value of Ω (see [2] or [7] for details). Typically, the same symmetric filter mask is used for all i.

If the number of constraint vectors n is larger than 3, T may be found via least-squares to be the eigenvector corresponding to the minimum eigenvalue of the matrix D where:

$$\mathrm{D} = \sum_{i=1}^{n} \tau_i \tau_i^{\mathrm{T}} \tag{5}$$

When noisy optical flow measurements were used, the recovered translation directions were noticeably biased towards the optical axis. Jepson and Heeger [7] explained this by illustrating that the covariance matrix $\Sigma_{\tilde{\tau}_i}$ for a noisy constraint $\tilde{\tau}_i$ represented a flattened ellipsoid rather than a sphere. The covariance matrix for a noisy twisted flow vector is:

$$\Sigma_{\tilde{q}_i} = \begin{bmatrix} (p_{xx})_i & (p_{xy})_i & (p_{xz})_i \\ (p_{xy})_i & (p_{yy})_i & (p_{yz})_i \\ (p_{xz})_i & (p_{yz})_i & (p_{zz})_i \end{bmatrix} = \sigma_n^2 \begin{bmatrix} f^2 & 0 & -fx_i \\ 0 & f^2 & -fy_i \\ -fx_i & -fy_i & (x_i)^2 + (y_i)^2 \end{bmatrix} \quad (6)$$

If (x_i, y_i) are the image plane coordinates of the center of the ith patch (i.e., the coordinates of $\tilde{\tau}_i$) and a symmetric convolution mask is used, then the relevant covariance matrix is:

$$\Sigma_{\tilde{\tau}_i} = \begin{bmatrix} (\tilde{\sigma}_{xx})_i & (\tilde{\sigma}_{xy})_i & (\tilde{\sigma}_{xz})_i \\ (\tilde{\sigma}_{xy})_i & (\tilde{\sigma}_{yy})_i & (\tilde{\sigma}_{yz})_i \\ (\tilde{\sigma}_{xz})_i & (\tilde{\sigma}_{yz})_i & (\tilde{\sigma}_{zz})_i \end{bmatrix} = \sigma_n^2 \begin{bmatrix} f^2 & 0 & -fx_i \\ 0 & f^2 & -fy_i \\ -fx_i & -fy_i & \alpha_i + \beta_i \end{bmatrix}$$

$$\alpha_i = (x_i)^2 + (y_i)^2 \qquad \beta_i = \sum_{j=1}^{n} (c_{ij})^2 \left[(x_j - x_i)^2 + (y_j - y_i)^2 \right] \quad (7)$$

If the same filter coefficients are used for calculating each τ_i, β_i will be a constant.

2.3 Dithering Method (Jepson and Heeger)

The covariance matrix in (7) represents a flattened ellipsoid which is much narrower in the mean sampling direction $\mathbf{s}_i$ than at right angles to $\mathbf{s}_i$. Jepson and Heeger [7] reasoned that by adding extra noise or *dithering*, the covariance matrix could be altered to approximate a sphere and this would eliminate the observed bias. This noise is added only along $\mathbf{s}_i$ since it is this dimension of the ellipsoid which is much smaller than the other two.

The expected value of the dithered least-squares matrix can be approximated as:

$$E[\bar{D}] \approx D + \kappa I_3 \quad (8)$$

where the value of κ depends upon the amount of noise. If $\{\lambda_1, \lambda_2, \lambda_3\}$ are the eigenvalues of D, $\{\lambda_1 + \kappa, \lambda_2 + \kappa, \lambda_3 + \kappa\}$ will be the eigenvalues of $\bar{D}$. Thus, the relative ordering of the eigenvalues will not be changed. Both D and $E[\bar{D}]$ will have the same eigenvectors.

2.4 Unbiasing Algorithm (Kanatani)

The expected value of the noisy least-squares matrix $\tilde{D}$ can be shown to be [2], [9, eq. 56]:

$$E[\tilde{D}] = D + \sum_i \Sigma_{\tilde{\tau}_i} \quad (9)$$

The $\Sigma_{\tilde{\tau}_i}$ terms are the covariance matrices of the constraint vectors as given in (7).

Clearly, $\tilde{D}$ provides a *biased* estimate of D. Now consider:

$$D' = \tilde{D} - \sum_i \Sigma_{\tilde{\tau}_i} \quad (10)$$

which is an unbiased estimate of D. $\tilde{D}$ can be obtained from the observed optical flow, and the $\Sigma_{\tilde{\tau}_i}$ terms can be calculated from known or estimated quantities. Thus, D' may be easily computed and used to generate an estimate for T. This operation was proposed by Kanatani [8, eq. 9] to obtain unbiased estimates in a least-squares situation.

2.5 Whitening Method (MacLean and Jepson)

The main disadvantage of the above two algorithms is their requirement for an accurate estimate of σ_n. To overcome this, MacLean and Jepson [10] defined:

$$M \equiv \frac{1}{\sigma_n^2} \sum_i \Sigma_{\tilde{\tau}_i} \quad (11)$$

Let $M^{1/2}$ be the matrix square root of M. From (9) and (11):

$$E\left[M^{-1/2}\tilde{D}M^{-1/2}\right] = M^{-1/2}\left(D + \sigma_n^2 M\right)M^{-1/2}$$
$$= M^{-1/2} D M^{-1/2} + \sigma_n^2 I \equiv E[\breve{D}] \quad (12)$$

Since D is symmetric, adding a multiple of the identity matrix will not affect its eigenvectors or the ordering of its eigenvalues. Consequently, the eigenvectors of $E[\breve{D}]$ are independent of the noise variance so σ_n^2 need not be estimated. If $\breve{T}$ is the minimum eigenvalue eigenvector of $\breve{D}$, the translation estimate in the original coordinate frame may be calculated as:

$$\hat{T} = M^{-1/2}\breve{T} \quad (13)$$

Note that $\hat{T}$ must be renormalized or scaled to have unit magnitude since $M^{-1/2}$ is not a length-preserving transformation.

2.6 Rotation and Depth Calculation

If $\mathbf{p}_i$ is a unit vector chosen perpendicular to the vector $A_i\hat{T}$ in (1) where $\hat{T}$ is the translation estimate, a least-squares estimate for the camera's rotational velocity may be obtained [4]:

$$\hat{\Omega} = \left[\sum_i B_i^T \mathbf{p}_i \mathbf{p}_i^T B_i \right]^{-1} \left[\sum_i B_i^T \mathbf{p}_i \mathbf{p}_i^T \mathbf{u}_i \right] \quad (14)$$

Once the quantities $\hat{T}$ and $\hat{\Omega}$ have been estimated, a relative depth map can be reconstructed from the optical flow data using least-squares:

$$\hat{Z}_i = \frac{(A_i\hat{T})^T (A_i\hat{T})}{(A_i\hat{T})^T (\mathbf{u}_i - B_i\hat{\Omega})} \quad (15)$$

2.7 Estimation Optimization

During simulations, it was observed that the $\hat{T}$ estimates were not always particularly accurate. Conversely, the maximum eigenvalue eigenvectors $\mathbf{v}_3$ were always almost exactly perpendicular to the true translation direction. Thus, it seemed as though the linear constraint technique could identify the plane containing T, but not necessarily T itself. Jepson and Heeger's method seems to break down when using a narrow viewing angle [2].

The original motivation for developing the linear constraint algorithm was to bypass the need for iteratively optimizing an error function over the entire possible range for $\hat{T}$ (i.e., the unit sphere). However, since $\mathbf{v}_3$ is known to be almost exactly perpendicular to T, it is possible to constrain the translation direction estimate to lie in the plane defined by $\mathbf{v}_1$ and $\mathbf{v}_2$. In fact, $\hat{T}$ can be written in terms of a single unknown rotation angle ϕ around $\mathbf{v}_3$.

$$\hat{T} = \cos(\phi)\mathbf{v}_1 + \sin(\phi)\mathbf{v}_2 \quad (16)$$

We can determine an appropriate value for ϕ by using an iterative approach to minimize the summed squared differences between the observed and expected optical flow.

3 Error Analysis

3.1 Least-Squares Matrix—Unbiasing Method

The means of the least-squares matrices $\tilde{D}$ and D' have already been calculated.

$$E[\tilde{D}] = D + \sum_i \Sigma_{\tilde{\tau}_i} \qquad E[D'] = D \quad (17)$$

Now suppose $\tilde{D}$ is rewritten as a vector $\tilde{\mathbf{d}}$ with the columns of $\tilde{D}$

being stacked vertically. It is then possible to calculate a covariance matrix for $\tilde{\mathbf{d}}$.

$$\Sigma_{\tilde{\mathbf{d}}} = \left[\tilde{\sigma}_{i,j}\right] = E\left[\tilde{\mathbf{d}}\tilde{\mathbf{d}}^{\mathrm{T}}\right] - E\left[\tilde{\mathbf{d}}\right]E\left[\tilde{\mathbf{d}}^{\mathrm{T}}\right] \tag{18}$$

For illustrative purposes, we shall only consider one entry in the covariance matrix—the remaining entries may be calculated in a similar fashion. Letting $\tau_i = \begin{bmatrix} x_i & y_i & z_i \end{bmatrix}^{\mathrm{T}}$ and $\Delta\tau_i = \begin{bmatrix} \delta x_i & \delta y_i & \delta z_i \end{bmatrix}^{\mathrm{T}}$, it can be shown that [2]:

$$\begin{aligned}\tilde{\sigma}_{2,6} &= E\left[\tilde{d}_2\tilde{d}_6\right] - E\left[\tilde{d}_2\right]E\left[\tilde{d}_6\right] \\ &= \sum_i\sum_j x_i y_j E\left[\delta y_i \delta z_j\right] + x_i z_j E\left[\delta y_i \delta y_j\right] + y_i y_j E\left[\delta x_i \delta z_j\right] + y_i z_j E\left[\delta x_i \delta y_j\right] \\ &+ \left\{\sum_k c_{ik}c_{jk}\left(p_{xy}\right)_k\right\}\left\{\sum_m c_{im}c_{jm}\left(p_{yz}\right)_m\right\} + \left\{\sum_k c_{ik}c_{jk}\left(p_{xz}\right)_k\right\}\left\{\sum_m c_{im}c_{jm}\left(p_{yy}\right)_m\right\}\end{aligned} \tag{19}$$

Since the $\tilde{\tau}$ vectors are calculated as a weighted local sum of twisted flow vectors, there will be dependency between $\tilde{\tau}_i$ and τ_j if their convolution masks overlap. In such a situation, the resulting "cross-covariance" can be calculated with the following formula [2]:

$$E\left[\delta x_i \delta y_j\right] = \sum_k c_{ik}c_{jk}\left(p_{xy}\right)_k \tag{20}$$

where the p_{xy} terms are defined in (6).

For the unbiased vector $\acute{\mathbf{d}}$, the covariance matrix $\Sigma_{\acute{\mathbf{d}}}$ will be the same as $\Sigma_{\tilde{\mathbf{d}}}$ since subtracting a constant vector simply shifts the mean and does not affect the covariance.

3.2 Least-Squares Matrix—Dithering Method

The mean of the dithered least-squares matrix $\overline{\mathrm{D}}$ can be shown to be [2]:

$$E\left[\overline{\mathrm{D}}\right] = \mathrm{D} + \sum_i\left\{\Sigma_{\tilde{\tau}_i} + \Sigma_{\overline{\tau}_i}\right\} \tag{21}$$

$\Sigma_{\tilde{\tau}_i}$ represents the covariance matrix of the undithered $\tilde{\tau}_i$ vector as given in (7). $\Sigma_{\overline{\tau}_i}$ is the covariance matrix of the dithered component which may be easily derived from the standard deviation of the dithering noise for $\tilde{\tau}_i(\sigma_{d,i}^2)$ and the direction in which it is applied.

$$\Sigma_{\overline{\tau}_i} = \begin{bmatrix} \left(\overline{\sigma}_{xx}\right)_i & \left(\overline{\sigma}_{xy}\right)_i & \left(\overline{\sigma}_{xz}\right)_i \\ \left(\overline{\sigma}_{xy}\right)_i & \left(\overline{\sigma}_{yy}\right)_i & \left(\overline{\sigma}_{yz}\right)_i \\ \left(\overline{\sigma}_{xz}\right)_i & \left(\overline{\sigma}_{yz}\right)_i & \left(\overline{\sigma}_{zz}\right)_i \end{bmatrix} = \frac{\sigma_{d,i}^2}{x_i^2 + y_i^2 + f^2}\begin{bmatrix} x_i^2 & x_i y_i & x_i f \\ x_i y_i & y_i^2 & y_i f \\ x_i f & y_i f & f^2 \end{bmatrix} \tag{22}$$

Noting that $\Delta\tau_i$ and $\Delta\overline{\tau}_i$ are mutually independent and that the individual dithering terms are independent, the sample covariance matrix entry can be shown to be:

$$\begin{aligned}\overline{\sigma}_{2,6} = \tilde{\sigma}_{2,6} + \sum_k \Big\{ & x_k y_k\left(\overline{\sigma}_{yz}\right)_k + x_k z_k\left(\overline{\sigma}_{yy}\right)_k + y_k y_k\left(\overline{\sigma}_{xz}\right)_k + y_k z_k\left(\overline{\sigma}_{xy}\right)_k \\ & + \left(\tilde{\sigma}_{xy}\right)_k\left(\overline{\sigma}_{yz}\right)_k + \left(\tilde{\sigma}_{xz}\right)_k\left(\overline{\sigma}_{yy}\right)_k + \left(\tilde{\sigma}_{yy}\right)_k\left(\overline{\sigma}_{xz}\right)_k + \left(\tilde{\sigma}_{yz}\right)_k\left(\overline{\sigma}_{xy}\right)_k \\ & + \left(\overline{\sigma}_{xy}\right)_k\left(\overline{\sigma}_{yz}\right)_k + \left(\overline{\sigma}_{xz}\right)_k\left(\overline{\sigma}_{yy}\right)_k \Big\}\end{aligned} \tag{23}$$

3.3 Least-Squares Matrix—Whitening Method

Let m_{kl} represent the (k, l) entry of the symmetric matrix $\mathrm{M}^{-1/2}$. It can easily be shown that entry (i, j) of $\breve{\mathrm{D}}$ may be calculated as:

$$\breve{d}_{ij} = \sum_{k=1}^{3}\sum_{l=1}^{3} m_{il}m_{kj}\tilde{d}_{lk} \tag{24}$$

If the least-squares matrix is converted to vector form, this can be written as:

$$\breve{\mathbf{d}} = \mathrm{N}\tilde{\mathbf{d}} \tag{25}$$

where N is a symmetric 9×9 matrix defined as:

$$\begin{aligned} &\mathrm{N} \equiv \left[n_{ij}\right] = \left[m_{pq}m_{rt}\right] \\ &p = \left((i-1)\bmod 3\right) + 1 \qquad r = \mathrm{floor}\left\{(j+2)/3\right\} \\ &q = \left((j-1)\bmod 3\right) + 1 \qquad t = \mathrm{floor}\left\{(i+2)/3\right\}\end{aligned} \tag{26}$$

The covariance matrix of $\breve{\mathbf{d}}$ may now be found to be:

$$\Sigma_{\breve{\mathbf{d}}} = E\left[\breve{\mathbf{d}}\breve{\mathbf{d}}^{\mathrm{T}}\right] - E\left[\breve{\mathbf{d}}\right]E\left[\breve{\mathbf{d}}^{\mathrm{T}}\right] = \mathrm{N}E\left[\tilde{\mathbf{d}}\tilde{\mathbf{d}}^{\mathrm{T}}\right]\mathrm{N}^{\mathrm{T}} - \mathrm{N}E\left[\tilde{\mathbf{d}}^{\mathrm{T}}\right]E\left[\tilde{\mathbf{d}}^{\mathrm{T}}\right]\mathrm{N}^{\mathrm{T}} = \mathrm{N}\Sigma_{\tilde{\mathbf{d}}}\mathrm{N}^{\mathrm{T}} \tag{27}$$

3.4 Translation Direction—Covariance Matrix

It is possible to approximate the covariance matrices for the eigenvectors of $\tilde{\mathrm{D}}$ using matrix perturbation theory [15]. Weng et al. [14] used a Taylor series expansion to derive the first-order perturbation terms for the eigenvectors as:

$$\Delta\mathbf{v}_i \approx \delta\mathbf{v}_i = \mathrm{V}\Delta_i\mathrm{V}^{\mathrm{T}}\Delta_{\tilde{\mathrm{D}}}\mathbf{v}_i \tag{28}$$

where the columns of V are the eigenvectors of the true D, $\Delta_{\tilde{\mathrm{D}}}$ represents the perturbation in the $\tilde{\mathrm{D}}$ matrix, and Δ_1 is defined as (Δ_2 and Δ_3 are defined similarly):

$$\Delta_1 \equiv \mathrm{diag}\left\{0, \left(\lambda_1 - \lambda_2\right)^{-1}, \left(\lambda_1 - \lambda_3\right)^{-1}\right\} \tag{29}$$

Note that $\Delta_{\tilde{\mathrm{D}}}\mathbf{v}_i$ may be equivalently rewritten as:

$$\Delta_{\tilde{\mathrm{D}}}\mathbf{v}_i = \begin{bmatrix} \left(v_x\right)_i \mathrm{I}_3 & \left(v_y\right)_i \mathrm{I}_3 & \left(v_z\right)_i \mathrm{I}_3 \end{bmatrix}\Delta_{\tilde{\mathbf{d}}} \equiv \mathrm{U}_i\Delta_{\tilde{\mathbf{d}}} \tag{30}$$

where $\Delta_{\tilde{\mathbf{d}}}$ is the perturbation in $\tilde{\mathbf{d}}$. The first-order eigenvector perturbations are zero-mean. The covariance matrix for $\tilde{\mathbf{v}}_i$ may therefore be approximated as [2]:

$$\Sigma_{\tilde{\mathbf{v}}_i} \approx \mathrm{V}\Delta_i\mathrm{V}^{\mathrm{T}}\mathrm{U}_i\Sigma_{\tilde{\mathbf{d}}}\mathrm{U}_i^{\mathrm{T}}\mathrm{V}\Delta_i\mathrm{V}^{\mathrm{T}} \tag{31}$$

The corresponding covariance matrices for the unbiasing and dithering approaches may be found by substituting $\Sigma_{\acute{\mathbf{d}}}$ or $\Sigma_{\overline{\mathbf{d}}}$, respectively, for $\Sigma_{\tilde{\mathbf{d}}}$. Unfortunately, since the T estimate must be renormalized after whitening, the approximation is not applicable for that situation.

3.5 Translation Direction—Mean Vector

During simulations, it was observed that the translation estimates were still somewhat biased, despite using unbiased estimates of the least-squares matrix. It is simple to show that $\hat{\mathrm{T}}$ is asymptotically unbiased as the number of τ_i vectors becomes large. However, an unbiased estimate of D does not necessarily imply that the corresponding estimate of T is unbiased.

3.5.1 *Second-Order Eigenvector Perturbations*

It can be shown [2] that extending the method of Weng et al. [14] leads to the expected values of the second-order perturbations of the eigenvector $\mathbf{v}_1 = \begin{bmatrix} v_x & v_y & v_z \end{bmatrix}^{\mathrm{T}}$

$$E\left[\delta^2\mathbf{v}_1\right] = \mathrm{G}_1\mathrm{E}_1\mathbf{v}_1 - \mathrm{G}_1\mathrm{G}_1\mathrm{F}_1\mathbf{v}_1 \tag{32}$$

where $\mathrm{G}_1 \equiv \mathrm{V}\Delta_1\mathrm{V}^{\mathrm{T}}$ is defined in terms of quantities specified in Section 3.4, E_1 is a symmetric 3×3 matrix with the $\sigma_{i,j}$s corresponding to the antibiasing technique used:

$$\begin{aligned}
e_{1,1} &= g_{1,1}\sigma_{1,1} + g_{2,2}\sigma_{2,2} + g_{3,3}\sigma_{3,3} + 2g_{1,2}\sigma_{1,2} + 2g_{1,3}\sigma_{1,3} + 2g_{2,3}\sigma_{2,3} \\
e_{1,2} &= g_{1,1}\sigma_{1,2} + g_{1,2}\sigma_{2,2} + g_{1,3}\sigma_{2,3} + g_{2,1}\sigma_{1,5} + g_{2,2}\sigma_{2,5} + g_{2,3}\sigma_{3,5} \\
&\quad + g_{3,1}\sigma_{1,6} + g_{3,2}\sigma_{2,6} + g_{3,3}\sigma_{3,6} \\
e_{1,3} &= g_{1,1}\sigma_{1,3} + g_{1,2}\sigma_{2,3} + g_{1,3}\sigma_{3,3} + g_{2,1}\sigma_{1,6} + g_{2,2}\sigma_{2,6} + g_{2,3}\sigma_{3,6} \\
&\quad + g_{3,1}\sigma_{1,9} + g_{3,2}\sigma_{2,9} + g_{3,3}\sigma_{3,9} \\
e_{2,2} &= g_{1,1}\sigma_{2,2} + g_{2,2}\sigma_{5,5} + g_{3,3}\sigma_{6,6} + 2g_{1,2}\sigma_{2,5} + 2g_{1,3}\sigma_{2,6} + 2g_{2,3}\sigma_{5,6} \\
e_{2,3} &= g_{1,1}\sigma_{2,3} + g_{1,2}\sigma_{5,3} + g_{1,3}\sigma_{6,3} + g_{2,1}\sigma_{2,6} + g_{2,2}\sigma_{5,6} + g_{2,3}\sigma_{6,6} \\
&\quad + g_{3,1}\sigma_{2,9} + g_{3,2}\sigma_{5,9} + g_{3,3}\sigma_{6,9} \\
e_{3,3} &= g_{1,1}\sigma_{3,3} + g_{2,2}\sigma_{6,6} + g_{3,3}\sigma_{9,9} + 2g_{1,2}\sigma_{3,6} + 2g_{1,3}\sigma_{3,9} + 2g_{2,3}\sigma_{6,9}
\end{aligned} \tag{33}$$

and F_1 is the symmetric 3×3 matrix:

$$\begin{aligned}
f_{1,1} &= v_x^2\sigma_{1,1} + v_y^2\sigma_{1,5} + v_z^2\sigma_{1,9} + 2v_xv_y\sigma_{1,2} + 2v_xv_z\sigma_{1,3} + 2v_yv_z\sigma_{1,6} \\
f_{1,2} &= v_x^2\sigma_{2,1} + v_y^2\sigma_{2,5} + v_z^2\sigma_{2,9} + 2v_xv_y\sigma_{2,2} + 2v_xv_z\sigma_{2,3} + 2v_yv_z\sigma_{2,6} \\
f_{1,3} &= v_x^2\sigma_{3,1} + v_y^2\sigma_{3,5} + v_z^2\sigma_{3,9} + 2v_xv_y\sigma_{3,2} + 2v_xv_z\sigma_{3,3} + 2v_yv_z\sigma_{3,6} \\
f_{2,2} &= v_x^2\sigma_{5,1} + v_y^2\sigma_{5,5} + v_z^2\sigma_{5,9} + 2v_xv_y\sigma_{5,2} + 2v_xv_z\sigma_{5,3} + 2v_yv_z\sigma_{5,6} \\
f_{2,3} &= v_x^2\sigma_{6,1} + v_y^2\sigma_{6,5} + v_z^2\sigma_{6,9} + 2v_xv_y\sigma_{6,2} + 2v_xv_z\sigma_{6,3} + 2v_yv_z\sigma_{6,6} \\
f_{3,3} &= v_x^2\sigma_{9,1} + v_y^2\sigma_{9,5} + v_z^2\sigma_{9,9} + 2v_xv_y\sigma_{9,2} + 2v_xv_z\sigma_{9,3} + 2v_yv_z\sigma_{9,6}
\end{aligned} \tag{34}$$

The first-order eigenvector perturbations are linear functions of the perturbation in the least-squares vector **d**. Since this latter quantity is zero-mean, the expected values of the first-order perturbations will also be zero. However, the expected values of the second-order expansion terms are clearly nonzero. The biased mean of the recovered translation direction vector can thus be approximated reasonably accurately as (in unnormalized form):

$$E\left[\hat{\mathrm{T}}\right] \approx \mathbf{v}_1 + E\left[\delta^2\mathbf{v}_1\right] \tag{35}$$

The first term on the right-hand side represents the desired value, whereas the second term represents an unwanted bias. Consequently, an unbiased estimate of the least-squares matrix D does not necessarily imply an unbiased estimate of the translation direction T.

3.5.2 Whitening Renormalization

From Section 2.5, it is necessary to renormalize the whitened translation direction estimates after transforming back to the original coordinate space. Since the lengths of the resulting vectors depend on their orientations, each vector must be scaled by a different value to have unit magnitude. Thus, the estimates are not weighted equally when the mean is calculated. The result is that the renormalization process introduces yet another bias.

Due to the nature of the renormalization equation, it is not possible to derive an analytical estimate of this bias. However, numerical estimation can be used to yield a value which agrees reasonably well with simulation data. By sampling a large number of evenly distributed candidate vectors over the unit sphere in whitened space, transforming each back to original space, renormalizing, and finally taking their average with each vector weighted by its corresponding pdf value, we can generate a suitable estimate of the biased mean.

3.6 Cramér-Rao Lower Bound

Since the T estimates are constrained to have unit magnitude, it is necessary to modify the standard CRLB as detailed by Gorman and Hero [3]. We have calculated the appropriate CRLB for our application [2], although the derivation is too lengthy to be reproduced here.

4 Experimental Results

4.1 Results From Simulated Data

The simulated camera had a focal length of 16 mm and an imaging array measuring 6.4 mm (or 64 pixels) on each side. The horizontal and vertical spacing (S) between adjacent τ_i vectors was varied between four and eight pixels to investigate the effects of reducing the number of constraint vectors. A 7×7 convolution mask was used to generate the τ_i vectors. Each simulation run consisted of 1,000 trials—this number was selected to ensure that sufficiently accurate sample statistics could be computed. The optical flow noise was randomly generated for each trial. The mean vector and covariance matrix for the translation direction estimates from each algorithm were determined, and the corresponding predicted values were calculated via the methods outlined in Section 3. The noise percentages tested were 5% and 10%.

The first data set used a translation of $\mathrm{T} = \begin{bmatrix}-10 & 0 & 20\end{bmatrix}$ (cm/s) normalized to $\mathrm{T} = \begin{bmatrix}-0.4472 & 0 & 0.8944\end{bmatrix}$ and a rotation of $\Omega = \begin{bmatrix}-0.05 & 0.0 & -0.1\end{bmatrix}$ (rad/s). Table 1 contains the mean translation estimates for the five methods with a spacing of eight pixels. Additional simulation results may be found in [2]. The error column contains the Euclidean distances between the tips of the estimated and ideal translation direction vectors. As expected, the biased estimates are noticeably biased towards the z axis. Unbiasing performs slightly worse than dithering for this particular example, whereas whitening is somewhat better. Optimizing provides the best estimate out of the five methods at lower noise levels.

Tables 2 and 3 show predicted (from Sections 3.4 and 3.5) and simulated mean vectors and covariance matrices for the translation estimates using 5% noise and a spacing of four pixels. All other parameters are the same as before. The predicted means are biased as discussed in Section 3.5. The constrained CRLB has been calculated assuming a biased estimator mean equal to that of the unbiasing technique. As a consequence, this bound is only approximate when applied to the other antibiasing algorithms. The unbiasing and whitening methods have smaller covariance matrices than the dithering method. Optimizing has the smallest covariance matrix, which is also close to the CRLB. In addition, it is the least biased of the five estimation algorithms. Thus, optimizing appears to produce the best results for this particular case, although it is computationally more expensive than the other techniques.

One problem noticed during experimentation was that the probability distributions of the two smallest least-squares eigenvalues often had significant overlap. This implied that in a number of trials the wrong eigenvector was selected as the initial value for $\hat{\mathrm{T}}$ in the iterative optimization search. This caused problems since only the local neighborhood of the initial $\hat{\mathrm{T}}$ value was considered as the allowable search space. By evaluating the error between the observed and expected optical flow using both $\mathbf{v}_1$ and $\mathbf{v}_2$ as estimates for T, the eigenvectors could be swapped if necessary to select the best initial $\hat{\mathrm{T}}$. The optimized translation estimates with eigenvector swapping in Table 1 showed a significant improvement.

4.2 Results From Real Data

The algorithms were applied to the synthetic Yosemite Valley image sequence [1] and a real sequence of moving columns [11]. In the former sequence, the flow fields were cropped to eliminate the moving clouds since their motion was independent of the camera translation.

The translation estimates are listed in Tables 4 and 5. The flow was estimated using Uras et al.'s [12] and Horn and Schunck's [5] methods, respectively. Since the correct flow fields were available, appropriate values for ρ (Section 2.1) could be obtained. When the true flow was used as noiseless input, ρ was set to zero and thus the biased, dithering, and unbiasing methods (B/D/U) were all identical. The optimized estimate (Section 2.7) was taken to be the true translation for the Yosemite sequence. The true translation for the columns sequence was calculated from the camera transfor-

TABLE 1
SAMPLE MEAN TRANSLATION ESTIMATES FOR SIMULATED OPTICAL FLOW (7 × 7 MASK, S = 8)

Method	Noise	# Swaps	Mean Translation Estimate			Error
Biased	5%	–	–.2963	.0045	.9551	0.1627
	10%	–	–.1371	.0201	.9904	0.3253
Dithered	5%	–	–.4270	–.0010	.9042	0.0224
	10%	–	–.2829	.0145	.9590	0.1771
Unbiased	5%	–	–.5074	–.0018	.8617	0.0685
	10%	–	–.1401	.0139	.9900	0.3220
Whitened	5%	–	–.4632	–.0005	.8863	0.0180
	10%	–	–.3665	–.0030	.9304	0.0884
Optimized	5%	–	–.4447	–.0001	.8957	0.0028
	10%	–	–.3073	.0102	.9516	0.1515
Optimized (Swapping)	5%	2	–.4478	–.0004	.8941	0.0008
	10%	313	–.4420	–.0043	.8970	0.0072

TABLE 2
PREDICTED AND SIMULATED TRANSLATION DIRECTION ESTIMATES (5% NOISE, S = 4)

Method	Pred/Sim	Mean Translation Estimate		
Ideal	P	– 0.4472	0.0000	0.8944
Biased	P	– 0.2739	0.0046	0.9617
	S	–0.2740	0.0045	0.9617
Dithered	P	–0.4431	–0.0000	0.8965
	S	–0.4491	–0.0013	0.8935
Unbiased	P	–0.4841	–0.0012	0.8750
	S	–0.4865	–0.0013	0.8737
Whitened	P	–0.4570	–0.0010	0.8895
	S	–0.4582	–0.0007	0.8889
Optimized	S	–0.4469	–0.0002	0.8946

TABLE 3
PREDICTED AND SIMULATED COVARIANCE MATRICES FOR T ESTIMATES (5% NOISE, S = 4)

Method	Covariance Matrix (× 10^{-2})					
	Predicted			Simulated		
	0.1470	0.0125	0.0404	0.1666	0.0120	0.0484
Biased	0.0125	0.0169	0.0034	0.0120	0.0180	0.0035
	0.0404	0.0034	0.0111	0.0484	0.0035	0.0143
	4.0900	0.0951	2.0291	4.8203	0.0836	1.8541
Dithered	0.0951	0.1604	0.0472	0.0836	0.1546	0.0543
	2.0291	0.0472	1.0067	1.8541	0.0543	1.3659
	1.0290	0.0624	0.5145	1.2491	1.0648	0.8102
Unbiased	0.0624	0.0390	0.0312	0.0648	0.0480	0.0427
	0.5145	0.0312	0.2573	0.8102	0.0427	0.5600
				0.5584	0.0588	0.3157
Whitened				0.588	0.0425	0.0326
				0.3157	0.0326	0.1835
				0.0044	0.0032	0.0022
Optimized				0.0032	0.0373	0.0016
				0.0022	0.0016	0.0011
	0.0290	0.0016	0.0161			
Constrained	0.0016	0.0243	0.0009			
CRLB	0.0161	0.0009	0.0089			

mation matrices for the images. The nonzero errors for the correct flow in Table 5 are due to inexact knowledge of the camera calibration parameters.

5 DISCUSSION

We have compared approaches for eliminating the bias in translation direction estimates obtained using linear constraints. Whitening produced the best results of the noniterative antibiasing methods. An error analysis indicated that dithering should result in the largest covariance matrix for the translation estimates. This prediction was supported by the simulation results. We introduced a new simple iterative search technique for optimizing the motion estimates. This method possessed the smallest covariance matrix of all five translation estimation methods and approached the constrained CRLB in simulation experiments.

An important result from the error analysis is that an unbiased estimate of a least-squares matrix does not necessarily yield unbiased eigenvalue and eigenvector estimates.

TABLE 4
RESULTS FOR YOSEMITE VALLEY SEQUENCE ($\rho = 0.148$, 31×31 MASK, S = 32)

Flow Type	Method	Translation Dir. Estimate			Error
Correct	B/D/U	−0.0041	0.3552	0.9348	0.0080
	Whitened	−0.0041	0.3552	0.9348	0.0080
	Optimized	0.0032	0.3582	0.9336	0.0000
Estimated	Biased	−0.1647	0.1387	0.9765	0.2797
	Dithered	−0.1780	0.1325	0.9751	0.2924
	Unbiased	−0.1639	0.1410	0.9763	0.2773
	Whitened	−0.0611	0.3992	0.9148	0.0785
	Optimized	−0.0829	0.3481	0.9338	0.0867

TABLE 5
RESULTS FOR COLUMNS SEQUENCE ($\rho = 0.314$, 11×11 MASK, S = 8)

Flow Type	Method	Translation Dir. Estimate			Error
Correct	B/D/U	−0.4756	−0.1427	0.8680	0.0270
	Whitened	−0.4756	−0.1427	0.8680	0.0270
	Optimized	−0.4752	−0.1426	0.8683	0.0265
Estimated	Biased	−0.2447	−0.0860	0.9658	0.2324
	Dithered	−0.3289	−0.1233	0.9363	0.1396
	Unbiased	−0.3088	−0.1112	0.9446	0.1621
	Whitened	−0.5024	−0.1878	0.8440	0.0836
	Optimized	−0.4599	−0.1708	0.8714	0.0439

ACKNOWLEDGMENTS

This research has been partially supported by grants from the Information Technology Research Centre of Ontario and the Natural Science and Engineering Research Council of Canada. This work was presented in part at the 1995 IEEE International Conference on Image Processing.

REFERENCES

[1] J. Barron, D. Fleet, and S. Beauchemin, "Performance of Optical Flow Techniques," *Int'l J. Computer Vision*, vol. 12, no. 1, pp. 43-77, 1994.

[2] A.M. Earnshaw and S.D. Blostein, "The Performance of Camera Translation Direction Estimators," Technical Report 95-10, Dept. of Electrical and Computer Eng., Queen's Univ., 1995.

[3] J.D. Gorman and A.O. Hero, "Lower Bounds for Parametric Estimation with Constraints," *IEEE Trans. Information Theory*, vol. 36, no. 6, pp. 1,285-1,301, 1990.

[4] D.J. Heeger and A.D. Jepson, "Subspace Methods for Recovering Rigid Motion I: Algorithm and Implementation," *Int'l J. Computer Vision*, vol. 7, no. 2, pp. 85-117, 1992.

[5] B.K. Horn and B.G. Schunck, "Determining Optical Flow," *Artificial Intelligence*, vol. 17, pp. 185-203, 1981.

[6] A.D. Jepson and D.J. Heeger, "A Fast Subspace Algorithm for Recovering Rigid Motion," *Proc. IEEE Workshop Visual Motion*, pp. 124-131, 1991.

[7] A.D. Jepson and D.J. Heeger, "Linear Subspace Methods for Recovering Translational Direction," *Spatial Vision in Humans and Robots*, L. Harris and M. Jenkin, eds., pp. 39-62. Cambridge Univ. Press, 1993.

[8] K. Kanatani, "Renormalization for Unbiased Estimation," *Proc. Fourth Int'l Conf. Computer Vision*, pp. 599-606, 1993.

[9] K. Kanatani, "Unbiased Estimation and Statistical Analysis of 3-D Rigid Motion from Two Views," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 1, pp. 37-50, Jan. 1993.

[10] W.J. MacLean, private communication, 1994.

[11] M. Otte and H.-H. Nagel, "Optical Flow Estimation: Advances and Comparisons," *Proc. Third European Conf. Computer Vision*, vol. I, pp. 61-60, 1994.

[12] S. Uras, F. Girosi, A. Verri, and V. Torre, "A Computational Approach to Motion Perception," *Biological Cybernetics*, vol. 60, no. 1, pp. 79-87, 1988.

[13] J. Weng, N. Ahuja, and T.S. Huang, "Optimal Motion and Structure Estimation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 9, pp. 864-884, Sept. 1993.

[14] J. Weng, T.S. Huang, and N. Ahuja, "Motion and Structure from Two Perspective Views: Algorithms, Error Analysis, and Error Estimation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 11, no. 5, pp. 451-476, May 1989.

[15] J. Wilkinson, *The Algebraic Eigenvalue Problem*. Oxford Univ. Press, 1965.