AT&T Technical Journal
Vol. 64, No. 2, February 1985
Printed in U.S.A.

Algorithms for Estimation of Three-Dimensional Motion

By A. N. NETRAVALI and J. SALZ*

(Manuscript received March 1, 1984)

We derive robust algorithms for estimating parameters of motion of rigid bodies that are observed by a television camera. Motion may be three-dimensional, containing both translational and rotational components, but the observations using the television camera are two-dimensional, i.e., projections on the camera plane. Our algorithms do not require a priori knowledge of any corresponding points in three- and two-dimensional spaces. We give both recursive as well as nonrecursive algorithms that minimize the error in intensity by using the estimated motion parameters. Our theory has applications in interframe coding, computer vision, and computer animation. The efficacy of our methods and the quality of the estimation procedures must await experimental verification.

I. INTRODUCTION

One of the most important problems in machine analysis of image sequences captured by a television camera is estimating the motion of objects in the field of view.[1] We have previously given algorithms for estimating the displacement vector when the motion is restricted to translation in a plane perpendicular to the camera axis.[2,3] This was later extended to situations where the illumination in the scene is spatially nonuniform[4] and to computationally more complex algorithms with better properties.[5] In this paper, we propose a further extension by developing algorithms for estimating parameters of three-

* Authors are employees of AT&T Bell Laboratories.

dimensional motion. The motion thus may have both the translational as well as rotational components, and the translation may not be in a plane perpendicular to the camera axis. Of course, although the object is three-dimensional and it moves in a three-dimensional space, the observations made by the television camera are still in a two-dimensional space, i.e., the object is observed by being projected from the object space to the image plane. Thus, information is lost in going from the three-dimensional object space to the two-dimensional image plane; this is a major source of difficulty in such estimation problems. It leads to nonunique solutions or ambiguous situations, unless additional information is made available. One such example of additional information is the correspondence of points in two- and three-dimensional space. This example is often used to determine camera position[6] or to make motion estimation unique.[7,8] However, in many practical problems such correspondence is either difficult or impossible to establish.

Our contribution in this paper is twofold. First, we develop equations of motion by noting the fact that a television camera creates a frame every thirtieth of a second. Most rigid body motion, in such a small amount of time, tends to be small. Therefore we develop models of incremental motion that each use three parameters for translation and rotation. Our second contribution is to give robust recursive and nonrecursive algorithms for estimating these parameters. The algorithms minimize the error in observed intensity by using these estimated motion parameters. Also, since the estimation algorithm is based on linearizing the intensity function, it is applicable in situations where the motion parameters are small. We also give an extension based on successive linearization that will work even when the motion is substantially large.

Some of the limitations of our approach should be pointed out. First, we are considering rigid body motion, i.e., no deformation of the body is allowed as a function of time. Second, parameters are estimated to minimize the intensity estimation error, and therefore, they may not correspond exactly to the true motion parameters, particularly since the problem may not have a unique solution due to loss of information in transforming from three-dimensionality to two-dimensionality. However, we believe that in most reasonable cases, parameters estimated by our procedure will be those corresponding to motion. Third, traditional difficulties with dynamic scene analysis, such as occlusion, spatial nonuniformity of motion parameters and illumination, and lack of proper segmentation, are largely ignored at this stage. They will be considered in our future work. Last, and perhaps most important, we have no simulation results to evaluate the performance of our algorithms. We hope, however, that since motion estimation is important in such diverse fields as computer animation, computer vision, and interframe coding, these algorithms will be specialized to many of these applications and then evaluated.

II. MOTION MODEL

In this section, we develop a model of three-dimensional motion that includes translation and rotation. The only constraint we impose is that the body in motion stay rigid. Let us assume that the location of different points changes in the object space as a result of object motion and that only a two-dimensional projection on the image plane is observable using a camera. (See Fig. 1.) Let a point P designated by a vector $\mathbf{r}$ = col. (x, y, z) move to another point P' designated by vector $\mathbf{r}'$ = col. (x', y', z'). Since the body stays rigid,

$$\bar{\mathbf{r}} = \tilde{\mathbf{R}}\bar{\mathbf{r}} + \bar{\mathbf{T}}, \tag{1}$$

where $\tilde{\mathbf{R}}$ is a three-by-three rotation matrix and $\bar{\mathbf{T}}$ is a three-dimensional translation vector. $\tilde{\mathbf{R}}$ can be represented in terms of the Eulerian

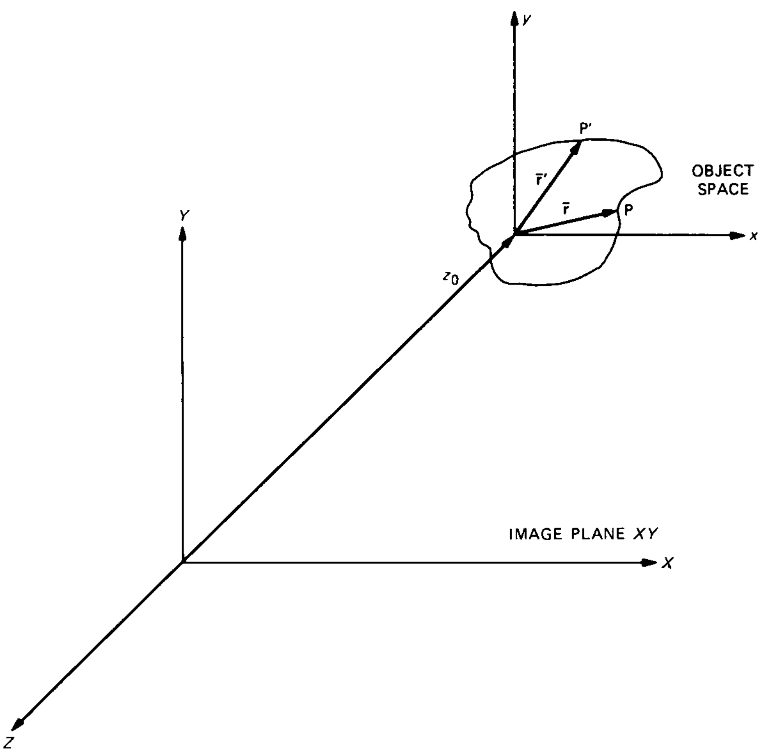

Fig. 1—Coordinate system showing object space and image plane.

angles ϕ, θ, and ψ as a product of three matrices, each corresponding to rotation about one axis. Thus

$$\tilde{\mathbf{R}} = \tilde{\mathbf{A}}\tilde{\mathbf{B}}\tilde{\mathbf{C}}, \tag{2}$$

where

$$\tilde{\mathbf{A}} = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3}$$

$$\tilde{\mathbf{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \tag{4}$$

$$\tilde{\mathbf{C}} = \begin{pmatrix} \cos\psi & 0 & -\sin\psi \\ 0 & 1 & 0 \\ \sin\psi & 0 & \cos\psi \end{pmatrix}. \tag{5}$$

These equations then specify general rigid body transformations. In practice, a television camera observes a new scene every thirtieth of a second. During such a small time, changes in the parameters of motion (i.e., θ, ϕ, ψ, and $\overline{\mathbf{T}}$) will be small. We therefore specialize these equations to small or infinitesimal changes in motion parameters that have taken place within a frame time.

2.1 Infinitesimal motion

For infinitesimal motion, changes in Euler angles are small. If these are denoted by $\Delta\theta$, $\Delta\phi$, and $\Delta\psi$, then if we use approximations, $\cos\Delta\theta = 1$ and $\sin\Delta\theta = \Delta\theta$, eqs. (3), (4), and (5) become

$$\tilde{\mathbf{A}} = \begin{pmatrix} 1 & \Delta\phi & 0 \\ -\Delta\phi & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{6}$$

$$\tilde{\mathbf{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \Delta\theta \\ 0 & -\Delta\theta & 1 \end{pmatrix} \tag{7}$$

$$\tilde{\mathbf{C}} = \begin{pmatrix} 1 & 0 & -\Delta\psi \\ 0 & 1 & 0 \\ \Delta\psi & 0 & 1 \end{pmatrix}. \tag{8}$$

Therefore,

$$\tilde{\mathbf{R}} = \tilde{\mathbf{A}}\tilde{\mathbf{B}}\tilde{\mathbf{C}}$$
$$\cong \begin{pmatrix} 1 & \Delta\phi & -\Delta\psi \\ -\Delta\phi & 1 & \Delta\theta \\ \Delta\psi & -\Delta\theta & 1 \end{pmatrix}$$
$$= \mathbf{I} + \begin{pmatrix} 0 & \omega_z & -\omega_y \\ -\omega_z & 0 & \omega_x \\ \omega_y & -\omega_x & 0 \end{pmatrix} \Delta t, \tag{9}$$

where $\mathbf{I}$ is the identity matrix; and ω_x, ω_y, and ω_z are angular velocities about the x, y, and z axes, respectively. By substituting (9) into (1), we get

$$\bar{\mathbf{r}}' = \bar{\mathbf{r}}_{t+\Delta t} = \bar{\mathbf{r}}_t + \tilde{\mathbf{P}}\bar{\mathbf{r}}_t\Delta t + \tilde{\mathbf{T}}\Delta \mathrm{t}, \tag{10}$$

where

$$\tilde{\mathbf{P}} = \begin{pmatrix} 0 & \omega_z & -\omega_y \\ -\omega_z & 0 & \omega_x \\ \omega_y & -\omega_x & 0 \end{pmatrix} \tag{11}$$

and

$$\tilde{\mathbf{T}} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}, \text{ vector of translational velocities.} \tag{12}$$

Next we denote the coordinates in the image plane by (X, Y). The transformation from object space to image plane then proceeds as follows:

$$X = \frac{z_0 x}{z} \tag{13}$$

$$Y = \frac{z_0 y}{z}, \tag{14}$$

where z_0 is the distance from the origin of the object space to the origin of the image plane (see Fig. 1). Now for small Δt, (10) becomes

$$x' = x + \omega_z y\Delta t - \omega_y z\Delta t + v_x\Delta t \tag{15}$$

$$y' = y - \omega_z x\Delta t + \omega_x z\Delta t + v_y\Delta t \tag{16}$$

$$z' = z + \omega_y x\Delta t - \omega_x y\Delta t + v_z\Delta t \tag{17}$$

and

$$\frac{x'}{z'} \cong \frac{x}{z} + \left(\omega_z \frac{y}{z} + \frac{v_x}{z} - \omega_y\right) \Delta t + \frac{x}{z}\left(-\omega_y \frac{x}{z} + \omega_x \frac{y}{z} - \frac{v_z}{z}\right) \Delta t. \tag{18}$$

Similarly,

$$\frac{y'}{z'} \cong \frac{y}{z} + \left(\omega_x + \frac{v_y}{z} - \omega_z \frac{x}{z}\right) \Delta t + \frac{y}{z}\left(-\omega_y \frac{x}{z} + \omega_x \frac{y}{z} - \frac{v_z}{z}\right) \Delta t. \quad (19)$$

By letting $z_0 = 1$ and using the definitions

$$X = \frac{x}{z}, \qquad X' = \frac{x'}{z'}$$

$$V_x = \frac{v_x}{z} \quad \text{and} \quad V_y = \frac{v_y}{z}, \quad (20)$$

we obtain

$$X' = X + (\omega_z Y - \omega_y + V_x)\Delta t + \left(-X^2\omega_y + XY\omega_x - X\frac{v_z}{z}\right) \Delta t \quad (21)$$

$$Y' = Y + (-\omega_z X + \omega_x + V_y)\Delta t + \left(-XY\omega_y + Y^2\omega_x - Y\frac{v_z}{z}\right) \Delta t. \quad (22)$$

Let $a = v_z/z$ be the magnification parameter; then the differential movement of the coordinates in the image plane for infinitesimal motion is as follows:

$$\frac{dX}{dt} = \frac{X' - X}{\Delta t} = \omega_z Y - \omega_y(1 + X^2) + \omega_x XY + V_x - aX \quad (23)$$

$$\frac{dY}{dt} = \frac{Y' - Y}{\Delta t} = -\omega_z X + \omega_x(1 + Y^2) - \omega_y XY + V_y - aY. \quad (24)$$

Thus there are six unknowns (ω_x, ω_y, ω_z, V_x, V_y, a) that need to be evaluated to quantify motion. The only values that can be observed are the intensities of the image in the present and the previous frames. Several techniques can be formulated to estimate these parameters. In the following, two techniques are described in detail.

III. MOTION ESTIMATION

The first technique deals with situations where the intensity changes only slightly as a result of small changes in motion parameters, whereas the second technique does successive linearization and therefore can handle larger changes in intensity. Let $\mathbf{I}(X, Y, t)$ be the intensity function at time t. Then differential changes in intensity are expected by

$$\mathbf{I}(X, Y, t + \Delta t) = \mathbf{I}(X + \Delta_x \Delta t, Y + \Delta_y \Delta t, t), \quad (25)$$

where

$$\Delta_x = \omega_z Y - \omega_y(1 + X^2) + \omega_x XY - aX + V_x \tag{26}$$

$$\Delta_y = -\omega_z X + \omega_x(1 + Y^2) - \omega_y XY - aY + V_y. \tag{27}$$

Expanding the intensity function in power series in Δt yields

$$\mathbf{I}(X, Y, t + \Delta t) = \mathbf{I}(X, Y, t) + \frac{\partial}{\partial X}\mathbf{I}(X, Y, t) \cdot [\Delta_x]\Delta t + \frac{\partial}{\partial Y}\mathbf{I}(X, Y, t)\cdot[\Delta_y]\Delta t. \tag{28}$$

Thus,

$$\begin{aligned}\frac{\mathbf{I}(X, Y, t + \Delta t) - \mathbf{I}(X, Y, t)}{\Delta t} &= \omega_z[\mathbf{I}_x(X, Y, t)Y - \mathbf{I}_y(X, Y, t)X] \\ &- \omega_y[\mathbf{I}_x(X, Y, t)(1 + X^2) + \mathbf{I}_y(X, Y, t)XY] \\ &+ \omega_x[\mathbf{I}_x(X, Y, t)XY + \mathbf{I}_y(X, Y, t)\cdot(1 + Y^2)] \\ &- a[\mathbf{I}_x(X, Y, t)X + \mathbf{I}_y(X, Y, t)Y] \\ &+ V_x[\mathbf{I}_x(X, Y, t)] + V_y[\mathbf{I}_y(X, Y, t)].\end{aligned} \tag{29}$$

Let (X_i, Y_i) be a pel deemed to be from the set of "moving-area" pels, i.e., the frame difference at these locations is above a certain prespecified threshold. Then, for each such moving-area pel define the following six-dimensional vector

$$\boldsymbol{\phi}_1 = \begin{bmatrix} \mathbf{I}_x(X_i, Y_i, t)Y_i - \mathbf{I}_y(X_i, Y_i, t)X_i \\ -\mathbf{I}_x(X_i, Y_i, t)(1 + X_i^2) - \mathbf{I}_y(X_i, Y_i, t)X_iY_i \\ \mathbf{I}_x(X_i, Y_i, t)X_iY_i + \mathbf{I}_y(X_i, Y_i, t)(1 + Y_i)^2 \\ -\mathbf{I}_x(X_i, Y_i, t)X_i - \mathbf{I}_y(X_i, Y_i, t)Y_i \\ \mathbf{I}_x(X_i, Y_i, t) \\ \mathbf{I}_y(X_i, Y_i, t) \end{bmatrix}. \tag{30}$$

If

$$\mathbf{C} = \text{col.}\ (\omega_z, \omega_y, \omega_x, a, V_x, V_y)$$

denotes the six-dimensional parameter vector that needs to be estimated, then we can express the measured intensity difference, M_i,

$$M_i = \boldsymbol{\phi}_1^T\mathbf{C} + \text{noise}. \tag{31}$$

If the number of measurements is n, then the problem is to create a least-squares estimate of $\mathbf{C}$ (labeled $\hat{\mathbf{C}}_n$) that minimizes the following Mean-Squared Error (MSE) after these n measurements:

$$\text{MSE} = \min_{\mathbf{C}_n} \sum_{i=1}^{n} (M_i - \boldsymbol{\phi}_1^T \mathbf{C}_n)^2. \tag{32}$$

Carrying out the minimization, we get the set of equations

$$\sum_{i=1}^{n} \boldsymbol{\phi}_1 M_i = \left[\sum_{i=1}^{n} \boldsymbol{\phi}_1 \boldsymbol{\phi}_1^T \right] \hat{\mathbf{C}}_n. \tag{33}$$

Thus, calculation of $\hat{\mathbf{C}}_n$ requires a matrix inversion at every step. The inversion can be carried out recursively as follows.
Let

$$\mathbf{A}_n = \sum_{i=1}^{n} \boldsymbol{\phi}_1 \boldsymbol{\phi}_1^T \tag{34}$$

and

$$\eta_n = \sum_{i=1}^{n} \boldsymbol{\phi}_1 M_i. \tag{35}$$

Clearly,

$$\boldsymbol{\eta}_n = \boldsymbol{\eta}_{n-1} + \boldsymbol{\phi}_n M_n \tag{36}$$

and

$$\mathbf{A}_n = \mathbf{A}_{n-1} + \boldsymbol{\phi}_n \boldsymbol{\phi}_n^T. \tag{37}$$

From the matrix inversion lemma in Ref. 9, we obtain

$$\mathbf{A}_n^{-1} = \mathbf{A}_{n-1}^{-1} - \frac{\mathbf{A}_{n-1}^{-1} \boldsymbol{\phi}_n \boldsymbol{\phi}_n^T \mathbf{A}_{n-1}^{-1}}{1 + \boldsymbol{\phi}_n^T \mathbf{A}_{n-1}^{-1} \boldsymbol{\phi}_n}, \tag{38}$$

and when this is used in (33), we get the recursion

$$\hat{\mathbf{C}}_n = \hat{\mathbf{C}}_{n-1} - \frac{\mathbf{A}_{n-1}^{-1}}{1 + \boldsymbol{\phi}_n^T \mathbf{A}_{n-1}^{-1} \boldsymbol{\phi}_n} \cdot \boldsymbol{\phi}_n (\boldsymbol{\phi}_n^T \hat{\mathbf{C}}_{n-1} - M_n). \tag{39}$$

If

$$\frac{\mathbf{A}_{n-1}^{-1}}{1 + \boldsymbol{\phi}_n^T \mathbf{A}_{n-1}^{-1} \boldsymbol{\phi}_n} = \alpha = \text{constant}, \tag{40}$$

then the above reduces to a simple gradient algorithm.

If the motion was purely translational in the image plane and there was no zooming (i.e., $v_z = 0$), then

$$\omega_z = \omega_x = \omega_y = a = 0. \tag{41}$$

The estimation of motion parameters would then be analogous to our previous schemes. Matrix $\mathbf{A}_n$ would be two-by-two, and the vectors such as $\boldsymbol{\eta}_n$, $\mathbf{C}_n$ would be two-dimensional.

3.1 Successive linearization estimation

In the previous section we derived an iterative procedure that did not use the previous estimates in the linearization process. Improvement may be obtained if the intensity function is linearized at different locations in the previous frame based on the value of the previous estimates.[2] Thus, as before, let $\hat{\mathbf{C}}_{n-1}$ be the estimate of six parameters of motion made after observing a patch of $(n-1)$ pels. We wish to revise this estimate to obtain $\hat{\mathbf{C}}_n$, which includes a patch of n pels obtained by adding a new pel to the previous patch of $(n-1)$ pels. Define

$$\hat{\Delta}_x^{n-1} = \hat{\omega}_z^{n-1} Y - \hat{\omega}_y^{n-1}(1 + X^2) + \hat{\omega}_x^{n-1} XY - \hat{a}^{n-1} X + \hat{V}_x^{n-1} \tag{42}$$

$$\hat{\Delta}_y^{n-1} = -\hat{\omega}_z^{n-1} X + \hat{\omega}_x^{n-1}(1 + Y^2) - \hat{\omega}_y^{n-1} XY - \hat{a}^{n-1} Y + \hat{V}_y^{n-1}. \tag{43}$$

Also, define a new cost function DFD$(\cdot)$ to be*

$$\begin{aligned}\mathrm{DFD}(X, Y, \hat{\Delta}_x^{n-1}, \hat{\Delta}_y^{n-1}, t) \\ = \mathbf{I}(X, Y, t + \Delta t) - \mathbf{I}(X + \hat{\Delta}_x^{n-1}\Delta t, Y + \hat{\Delta}_y^{n-1}\Delta t, t).\end{aligned} \tag{44}$$

We note that, as defined above, DFD = 0 if

$$\begin{aligned}\hat{\Delta}_x^{k-1} &= \Delta_x, \quad \text{for any} \quad X, Y \\ \hat{\Delta}_y^{k-1} &= \Delta_y, \quad \text{for any} \quad X, Y,\end{aligned} \tag{45}$$

i.e., when our estimates of motion are equal to the true value of the parameters of motion. Also, for any given estimate of the motion parameters, DFD can be calculated if we know the intensities of two successive frames. As before, we can now expand DFD in Taylor's series. Thus

$$\begin{aligned}&\mathrm{DFD}(X, Y, \hat{\Delta}_x^{n-1}, \hat{\Delta}_y^{n-1}, t) \\ &= \mathbf{I}(X, Y, t + \Delta t) - \mathbf{I}(X + \hat{\Delta}_x^{n-1}\Delta t, Y + \hat{\Delta}_y^{n-1}\Delta t, t) \\ &= \mathbf{I}(X + \Delta_x\Delta t, Y + \Delta y\Delta t, t) - \mathbf{I}(X + \hat{\Delta}_x^{n-1}\Delta t, Y + \hat{\Delta}_y^{n-1}\Delta t, t) \\ &= \mathbf{I}(X + \hat{\Delta}_x^{n-1}\Delta t + (\Delta_x - \hat{\Delta}_x^{n-1})\Delta t, Y + \hat{\Delta}_y^{n-1}\cdot\Delta t \\ &\quad + (\Delta_y - \hat{\Delta}_y^{n-1})\Delta t, t) - \mathbf{I}(X + \hat{\Delta}_x^{n-1}\cdot\Delta t, Y + \hat{\Delta}_y^{n-1}\Delta t, t).\end{aligned} \tag{46}$$

Let

$$\bar{\mathbf{I}} = \mathbf{I}(X + \hat{\Delta}_x^{n-1}\Delta t, Y + \hat{\Delta}_y^{n-1}\Delta t, t) \tag{47}$$

---

* As in Ref. 2, DFD stands for displaced frame difference.

$$\frac{\mathrm{DFD}(X,\ Y,\ \hat{\Delta}_x^{n-1},\ \hat{\Delta}_y^{n-1},\ t)}{\Delta t}$$

$$\begin{aligned}
&\cong (\omega_z - \hat{\omega}_z^{n-1})[\overline{\mathbf{I}}_x Y - \overline{\mathbf{I}}_y X] \\
&\quad - (\omega_y - \hat{\omega}_y^{n-1})[\overline{\mathbf{I}}_x(1 + X^2) + \overline{\mathbf{I}}_y XY] \\
&\quad + (\omega_x - \hat{\omega}_x^{n-1})[\overline{\mathbf{I}}_x XY + \overline{\mathbf{I}}_y(1 + Y^2)] \\
&\quad + (\hat{a}^{n-1} - a)[\overline{\mathbf{I}}_x X + \overline{\mathbf{I}}_y Y] \\
&\quad + (V_x - \hat{V}_x^{n-1})[\overline{\mathbf{I}}_x] \\
&\quad + (V_y - \hat{V}_y^{n-1})[\overline{\mathbf{I}}_y]. \qquad (48)
\end{aligned}$$

Once again, define a vector of measured quantities

$$\boldsymbol{\phi}_1^{n-1} = \begin{bmatrix} \overline{\mathbf{I}}_x Y_i - \overline{\mathbf{I}}_y X_i \\ -\overline{\mathbf{I}}_x(1 + X_i^2) - \overline{\mathbf{I}}_y X_i Y_i \\ \overline{\mathbf{I}}_x X_i Y_i + \overline{\mathbf{I}}_y(1 + Y_i^2) \\ -\overline{\mathbf{I}}_x X_i - \overline{\mathbf{I}}_y Y_i \\ \overline{\mathbf{I}}_x \\ \overline{\mathbf{I}}_y \end{bmatrix}. \qquad (49)$$

Then

$$\begin{aligned}
M_i^{n-1} &= \mathrm{DFD}(X,\ Y,\ \hat{\Delta}_x^{n-1},\ \hat{\Delta}_y^{n-1},\ t)/\Delta t \\
&= (\boldsymbol{\phi}_1^{n-1})^{\mathbf{T}}(\mathbf{C} - \hat{\mathbf{C}}_{n-1}) + \text{noise}. \qquad (50)
\end{aligned}$$

The least-squares estimate, $\hat{\mathbf{C}}_n$, should minimize the following mean-squared error,

$$\mathrm{MSE} = \min_{\mathbf{C}} \left\{ \sum_{i=1}^{n} [M_i^{n-1} - (\boldsymbol{\phi}_i^{n-1})^{\mathbf{T}}(\mathbf{C} - \hat{\mathbf{C}}_{n-1})]^2 \right. \qquad (51)$$

for any given initial estimate $\hat{\mathbf{C}}_{n-1}$. As before, carrying out the minimization, we get

$$\sum_{i=1}^{n} \boldsymbol{\phi}_i^{n-1} \cdot M_i^{n-1} = \left[\sum_{i=1}^{n} \boldsymbol{\phi}_i^{n-1} \boldsymbol{\phi}_i^{n-1^{\mathbf{T}}} \right] (\hat{\mathbf{C}}_n - \hat{\mathbf{C}}_{n-1}). \qquad (52)$$

Then

$$\hat{\mathbf{C}}_n = \hat{\mathbf{C}}_{n-1} + \left[\sum_{i=1}^{n} \boldsymbol{\phi}_i^{n-1} \boldsymbol{\phi}_i^{n-1^{\mathbf{T}}} \right]^{-1} \left[\sum_{i=1}^{n} \boldsymbol{\phi}_i^{n-1} M_i^{n-1} \right]. \qquad (53)$$

As in the previous section, the matrix inversion lemma can be used to invert the matrix

$$\left[\sum_{i=1}^{n} \boldsymbol{\phi}_i^{n-1} \boldsymbol{\phi}_i^{n-1^{\mathrm{T}}}\right].$$

The real difference between this method and that of the previous section is that even if the motion is large (i.e., parameters of motion are somewhat large), the successive linearization, if it converges, gives more accurate estimates, since $(\mathbf{C} - \hat{\mathbf{C}}_{n-1})$ becomes smaller as iterations proceed.

IV. CONCLUSIONS

A mathematical theory that provides algorithms for robust estimation of a general set of motion parameters from frame sequences obtained from a television camera is now available. The theory was derived under mild assumptions. Both recursive and nonrecursive algorithms are provided. The efficacy of our algorithms has to be evaluated for each application. Potential applications are for interframe coding, computer vision, and computer animation.

REFERENCES

1. H. H. Nagel, "Overview of Image Sequence Analysis," in *Image Sequence Processing and Dynamic Scene Analysis*, T. S. Huang, Ed., New York: Springer-Verlag, 1983.
2. A. N. Netravali and J. D. Robbins, "Motion Compensated Television Coding," B.S.T.J., *58*, No. 3 (March 1979), pp. 629–68.
3. J. A. Stuller and A. N. Netravali, "Transform Domain Motion Estimation," B.S.T.J., *58*, No. 7 (September 1979), pp. 1673–782.
4. J. A. Stuller, A. N. Netravali, and J. D. Robbins, "Interframe Television Coding Using Gain and Displacement Compensation," B.S.T.J., *59*, No. 7 (September 1980), pp. 1227–40.
5. A. N. Netravali and J. D. Robbins, "Motion Compensated Television Coding: Some New Results," B.S.T.J., *59*, No. 9 (November 1980), pp. 1735–45.
6. S. K. Ganapathy, unpublished work.
7. R. Tsai and T. S. Huang, "Uniqueness and Estimation of Three Dimensional Motion Parameters of Rigid Objects with Curved Surfaces," CSL Technical Report R-921, University of Illinois, 1981.
8. B. Yen and T. S. Huang, "Determination of 3-D-Motion and Structure of a Rigid Body Using Spherical Projection," CSL Technical Report, University of Illinois, 1982.
9. B. D. O. Anderson and J. B. Moore, *Optimal Filtering, Information and System Science Series*, T. Kailath, Ed., Englewood Cliffs, NJ: Prentice-Hall, 1979.

AUTHORS

Arun N. Netravali, B. Tech. (Honors), 1967, Indian Institute of Technology, Bombay, India; M.S., 1969, Ph.D. (Electrical Engineering), 1970, Rice University; Optimal Data Corporation, 1970–1972; AT&T Bell Laboratories, 1972—. Mr. Netravali has worked on problems related to filtering, guidance, and control for the space shuttle. At AT&T Bell Laboratories, he has worked on various aspects of digital processing and computing. He was a Visiting Professor in the Department of Electrical Engineering at Rutgers University. He is presently Director of the Computer Technology Research Laboratory.

Mr. Netravali holds over twenty patents and has had more than sixty papers published. He was the recipient of the Donald Fink Prize Award for the best review paper published in the Proceedings of the IEEE and the SMPTE journal award for the best paper published in 1982. Editorial board, Proceedings of the IEEE; Editor, IEEE Transactions on Communications; fellow, IEEE; member, Tau Beta Pi, Sigma Xi.

Jack Salz, B.S.E.E., 1955, M.S.E., 1956, and Ph.D., 1961, University of Florida; AT&T Bell Laboratories, 1961—. Mr. Salz first worked on the electronic switching system. Since 1968 he has supervised a group engaged in theoretical studies in data communications and is currently a member of the Communications Methods Research Department. During the academic year 1967–68, he was on leave as Professor of Electrical Engineering at the University of Florida. He was a visiting lecturer at Stanford University in Spring 1981 and a visiting MacKay Lecturer at the University of California, at Berkeley, in Spring 1983.