Hybrid Feature/Flow-Based 3-D Trajectory Estimation *

Steven D. Blostein and Robert M. Chann
Department of Electrical Engineering
Queen's University
Kingston, Ontario, K7L 3N6, CANADA

Abstract

This paper examines the problem of estimating three-dimensional trajectories of moving objects from noisy monocular image sequences. We have built on the basic models and formulation of Broida et al. [2] to estimate the motion of a rigid object moving in front of a stationary camera at constant translational and rotational velocities. The main emphasis of this paper is on the development of a hybrid feature/flow-based estimation procedure in which both feature point and optical flow measurements are utilized. We have also analyzed abrupt changes in object trajectories and how process noise can be used to improve the performance of a filter in such situations. Preliminary experiments have shown some very encouraging results.

1 Introduction

Recovering the three-dimensional (3-D) motion of rigid object(s) from image sequences is a fundamental problem in computer vision. There are two basic approaches to solving this problem, namely, the feature-based approach and the optical flow-based approach. In feature-based approaches, 3-D motion is derived by observing the 2-D positions of a set of relatively sparse discrete features. In contrast, optical flow-based techniques derive 3-D motion from relatively dense 2-D instantaneous velocity data.

The feature-based approach can be further divided into recursive formulations [2][5][7] and batch formulations [1][10]. In recursive formulations, the computation of a new estimate is based only on the current measurement(s) and estimate, thus making it suitable for real-time applications. Silvén and Repo [7] implemented a complete real-time tracking system based on the extended Kalman filter (EKF). Iu and Wohn [5] used the same filter but proposed a method for the estimation of motion modelled by an arbitrary number of derivatives. Broida *et al.* [2] formulated a recursive procedure using the iterated extended Kalman filter (IEKF) because it is less likely to diverge than the EKF.

In our work, we assume an environment where a stationary video camera is used to observe a single rigid moving object over time. Our goal is to recover the 3-D trajectory of this object from an extended image sequence captured by the video camera. This paper begins by examining the various models that are essential to our research, followed by a description of how optical flow measurements can be incorporated into the recursive formulation proposed by Broida *et al.* An extensive set of simulations has been designed and will be described in details. An analysis of the results and a discussion of issues that require further research will be presented. Finally, conclusions will be drawn based on the simulation results.

*This work is supported partially by NSERC and partially by IRIS under the NCE program of the Government of Canada.

2 Models

2.1 Imaging Model

The image acquisition process is modelled using the perspective projection. Consider a 3-D point $\mathbf{p} = (x, y, z)$. The projection of $\mathbf{p}$ on to the image plane through the perspective transformation is described by the vector equation

$$\mathbf{P} = \begin{bmatrix} X \\ Y \end{bmatrix} = \frac{f}{z} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} n_x \\ n_y \end{bmatrix} \qquad (1)$$

where n_x and n_y represent additive noise terms that are included to account for random errors such as those arising from spatial quantization. Without loss of generality, we shall assume the focal length to be unity (i.e., $f = 1$) throughout this paper.

2.2 Object Model

Our goal is to observe a moving object, whose structure is unknown, through a stationary camera; therefore, it is useful to define a camera-centred coordinate system (CCCS) with its $z-$axis pointing in the direction of the optical axis. An object-centred coordinate system (OCCS) is also needed to facilitate the estimation of the object structure. The following equation allows the conversion from the OCCS to the CCCS:

$$\mathbf{p}(t) = \mathbf{p}_R(t) + R(t)\mathbf{p}_O \qquad (2)$$

where $\mathbf{p}_O$ is a point in the OCCS, $\mathbf{p}(t)$ is the same point represented in the CCCS, $R(t)$ is the rotation matrix that aligns the OCCS with the CCCS and $\mathbf{p}_R(t)$ denotes the origin of the OCCS which is *not* directly observable. Due to the rigidity assumption, there is no time dependency on $\mathbf{p}_O$.

2.3 Motion Model

It is well-known that the motion of a rigid body is completely characterized by its translational and rotational motion [3]. The models for these two types of motion are stated below.

2.3.1 Translational Motion

A linear model is used to describe the translational motion. In other words, we assume constant translational velocity ($\mathbf{T} = [T_x\ T_y\ T_z]^T$). Consider a point on the surface of the translating object and its (initial) location at time $t = t_0$. Its position at any time t is simply described by

$$\mathbf{x}(t) = \mathbf{x}(t_0) + (t - t_0)\mathbf{T} \tag{3}$$

2.3.2 Rotational Motion

The unit quaternion is used to represent the orientation of a rotating object and by assuming constant rotational velocity ($\omega = [\omega_x\ \omega_y\ \omega_z]^T$), it can be propagated in time using

$$\mathbf{q}(t) = \exp[\Omega(t - t_0)]\mathbf{q}(t_0) \tag{4}$$

where

$$\Omega = \frac{1}{2}\begin{bmatrix} 0 & \omega_z & -\omega_y & \omega_x \\ -\omega_z & 0 & \omega_x & \omega_y \\ \omega_y & -\omega_x & 0 & \omega_z \\ -\omega_x & -\omega_y & -\omega_z & 0 \end{bmatrix} \tag{5}$$

In addition, the rotation matrix $R(t)$ in (2) can be expressed in terms of the unit quaternion (see [2]).

2.4 System Model

Our system consists of a rigid object translating and rotating continuously in time and observations that are made at regular time intervals. This is equivalent to a system having a continuous plant and discrete measurements [6]. The vector differential equation that describes the continuous plant is

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}(t), t) + G(t)\mathbf{w}(t) \tag{6}$$

where $\mathbf{x}(t)$ is the state of the system, $\mathbf{w}(t)$ represents the process noise and $G(t)$ is a matrix mapping $\mathbf{w}(t)$ into "state space". Observations, $\mathbf{y}_k$, are described by the discrete measurement equation

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k \tag{7}$$

where $\mathbf{h}(\cdot)$ is a nonlinear measurement function (due mainly to the perspective transformation) and $\mathbf{v}_k$ is the measurement noise. The problem is to find $\hat{\mathbf{x}}_k$, an estimate of $\mathbf{x}_k$, based on $\mathbf{y}_k$. Since both $\mathbf{f}(\mathbf{x}(t), t)$ and $\mathbf{h}(\mathbf{x}_k)$ are nonlinear, linear Kalman filter theory does not apply. A common approach is to use an approximate nonlinear filter such as the EKF instead. However, due to nonlinearities in $\mathbf{f}(\mathbf{x}(t), t)$ and the high degree of nonlinearity in $\mathbf{h}(\mathbf{x}_k)$, a more "sophisticated" filter, such as the IEKF or the iterated linear filter-smoother (ILFS), may result in improved performance. An overview of these three approximate nonlinear Kalman filters has been omitted due to space limitations but a detailed description can be found in [6].

All three filters mentioned above involve linearization of $\mathbf{f}(\mathbf{x}(t), t)$ and/or $\mathbf{h}(\mathbf{x}_k)$; therefore their derivatives

$$F(\hat{\mathbf{x}}(t), t) = \left.\frac{d\mathbf{f}(\mathbf{x}(t), t)}{d\mathbf{x}(t)}\right|_{\mathbf{x}(t)=\hat{\mathbf{x}}(t)} \tag{8}$$

and

$$H(\hat{\mathbf{x}}_k) = \left.\frac{d\mathbf{h}(\mathbf{x}_k)}{d\mathbf{x}_k}\right|_{\mathbf{x}_k=\hat{\mathbf{x}}_k} \tag{9}$$

are required. $\hat{\mathbf{x}}(t)$ and $\hat{\mathbf{x}}_k$ denote the points about which $\mathbf{f}(\mathbf{x}(t), t)$ and $\mathbf{h}(\mathbf{x}_k)$ are linearized, respectively.

As in [2], our state vector consists of the image plane coordinates of the origin of the OCCS, the "normalized" translational velocity, the unit quaternion representing the orientation of the object, the rotational velocity and the "normalized" coordinates of the feature points in the OCCS:

$$\mathbf{x}(t) = \begin{bmatrix} x_R(t)/z_R(t) \\ y_R(t)/z_R(t) \\ T_x/z_R(t) \\ T_y/z_R(t) \\ T_z/z_R(t) \\ q_1 \\ q_2 \\ q_3 \\ q_4 \\ \omega_x \\ \omega_y \\ \omega_z \\ x_1/z_R(t) \\ y_1/z_R(t) \\ z_1/z_R(t) \\ \vdots \\ x_M/z_R(t) \\ y_M/z_R(t) \\ z_M/z_R(t) \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \\ s_7 \\ s_8 \\ s_9 \\ s_{10} \\ s_{11} \\ s_{12} \\ s_{13} \\ s_{14} \\ s_{15} \\ \vdots \\ s_{10+3M} \\ s_{11+3M} \\ s_{12+3M} \end{bmatrix} \tag{10}$$

The normalization factor $z_R(t)$ is used to eliminate the unknown scaling factor due to the perspective transformation [2].

3 Combining Feature Point and Optical Flow Measurements

There has been no attempt to integrate the feature-based and optical flow-based approaches to 3-D motion estimation. The work presented in this section may well be the first step towards such a goal. We have included optical flow measurements, in addition to feature point position measurements, in our observations. By measuring the instantaneous 2-D velocities at or near pixel locations chosen as "features", we hope to improve the accuracy of the state estimates. The basic optical flow equations are [3]

$$u_i = \frac{T_x}{Z_i} - x_i\frac{T_z}{Z_i} - x_i y_i\omega_x + (1 + x_i^2)\omega_y - y_i\omega_z \tag{11}$$

$$v_i = \frac{T_y}{Z_i} - y_i\frac{T_z}{Z_i} - (1 + y_i^2)\omega_x + x_i y_i\omega_y + x_i\omega_z \tag{12}$$

or in the more compact vector-matrix form,

$$\mathbf{V}_i = \begin{bmatrix} u_i \\ v_i \end{bmatrix} = \frac{1}{Z}\begin{bmatrix} 1 & 0 & -x_i \\ 0 & 1 & -y_i \end{bmatrix}\begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} + \begin{bmatrix} -x_i y_i & (1+x_i^2) & -y_i \\ -(1+y_i^2) & x_i y_i & x_i \end{bmatrix}\begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \quad (13)$$

where a necessary sign reversal has been applied.

In order to incorporate optical flow data into our recursive filtering framework, (11) and (12) must be expressed in terms of the existing system states. First, we need to establish the following relationships

$$\frac{T_x}{Z_i} = \frac{T_x}{Z_R}\frac{Z_R}{Z_i} = \frac{s_3}{s_{12+3i}} \quad (14)$$

$$\frac{T_y}{Z_i} = \frac{T_y}{Z_R}\frac{Z_R}{Z_i} = \frac{s_4}{s_{12+3i}} \quad (15)$$

$$\frac{T_z}{Z_i} = \frac{T_z}{Z_R}\frac{Z_R}{Z_i} = \frac{s_5}{s_{12+3i}} \quad (16)$$

$$x_i = \frac{X_i}{Z_i} = \frac{X_i}{Z_R}\frac{Z_R}{Z_i} = \frac{s_{10+3i}}{s_{12+3i}} \quad (17)$$

$$y_i = \frac{Y_i}{Z_i} = \frac{Y_i}{Z_R}\frac{Z_R}{Z_i} = \frac{s_{11+3i}}{s_{12+3i}} \quad (18)$$

Then, we express (11) and (12) in terms of the system states defined in (10). We now have

$$u_i = \frac{s_3}{s_{12+3i}} - \frac{s_5 s_{10+3i}}{s_{12+3i}^2} - \frac{s_{10} s_{10+3i} s_{11+3i}}{s_{12+3i}^2} + (1 + \frac{s_{10+3i}^2}{s_{12+3i}^2})s_{11} - \frac{s_{12} s_{11+3i}}{s_{12+3i}} \quad (19)$$

$$v_i = \frac{s_4}{s_{12+3i}} - \frac{s_5 s_{11+3i}}{s_{12+3i}^2} - (1 + \frac{s_{11+3i}^2}{s_{12+3i}^2})s_{10} + \frac{s_{11} s_{10+3i} s_{11+3i}}{s_{12+3i}^2} + \frac{s_{12} s_{10+3i}}{s_{12+3i}} \quad (20)$$

where $i = 1 \ldots M$ with M denoting the total number of feature points. The matrix $H(\mathbf{x}_k)$ described in the previous section will have to be expanded as there are now 2 extra observations at each feature point location. To compute $H(\mathbf{x}_k)$, (19) and (20) must be differentiated with respect to each state.

4 Simulation Results

In order to study the benefits of including optical flow measurements as observations, simulations have been performed. We have designed a more comprehensive collection of experiments than those outlined in [2]. First of all, all simulations are carried out in Monte Carlo fashion. Secondly, the experiments are designed to evaluate the proposed method under a wide variety of conditions. Moreover, unlike most of the existing work in this area, *three* filters (the EKF, IEKF and ILFS) are considered because the relative performance of non-linear filters is highly dependent on the particular problem at hand [9].

Due to space limitation, the following strategy is used to reduce the number of graphs that have to be presented. Since we are only interested in the position $(1-2)$ and velocity $(3-5, 10-12)$ states, we omit the quaternion $(6-9)$ and structure $(13-24)$ states in our presentation, reducing the number of graphs to 8. Among these 8 remaining graphs, a representative one is chosen to show a position, translational velocity or rotational velocity state.

4.1 Implementation Details

Our "synthetic world" consists of a room of size $20 \times 25 \times 10$ cubic units, a stationary camera and a moving cube of size $3 \times 3 \times 3$ cubic units. Four of the corners of the cube are used as feature points, that is, $M = 4$. Each simulation is made up of 30 trial runs and each trial run consists of 100 image frames. Unless otherwise specified, the motion parameters used in this paper are $\mathbf{T} = [0.15\ 0.2\ 0.05]$ and $\omega = [0.05\ 0.05\ 0.05]$.

Errors in the initial state estimates are chosen to be between 20% and 40%. The initial quaternion states are always error-free since we assume that the OCCS is initially aligned with the CCCS. We pick the variance of the measurement noise to correspond to the spatial quantization noise found in a 512×512 image sequence where the object occupies an image area of approximately 200×200 pixels. For simplicity, the same amount of noise is added to both feature point and optical flow measurements.

4.2 Comparing the EKF, IEKF and ILFS

The aim of this experiment is to compare and evaluate the EKF, IEKF and ILFS for the purpose of 3-D trajectory estimation. Three factors should be taken into consideration. First of all, it is essential to determine which filter performs the "best" under a specific set of conditions. The best filter in each simulation is chosen as the filter that produces the smallest mean-squared error (MSE) the most number of times (with a maximum of 8 times). The MSE is defined as

$$MSE = \frac{1}{90}\sum_{i=11}^{100}(\bar{\hat{\mathbf{x}}}_i - \mathbf{x}_i)^2 \quad (21)$$

where $\mathbf{x}_i$ represents the true state vector in frame i and $\bar{\hat{\mathbf{x}}}_i$ is the Monte-Carlo-averaged estimate for the *i-th* frame, which is, in turn, defined as

$$\bar{\hat{\mathbf{x}}}_i = \frac{1}{30}\sum_{j=1}^{30}\hat{\mathbf{x}}_{ij} \quad (22)$$

The summation in (21) excludes the first 10 frames because wild fluctuations are likely during the early stage. For convenience, we use BC (or Best Count) to indicate the number of times a filter is chosen as the best.

Aside from determining the best among the three filters, it is equally important to decide if the filters,

	EKF	IEKF	ILFS
BC	25	7	0
PEA	94.2%	91.2%	90.8%

Table 1: Comparing the EKF, IEKF and ILFS

in fact, produce reasonable estimates. To achieve this goal, we first compute an average value for each of the 8 states, based on the actual trajectory in each experiment. If the MSE for a certain state is less than or equal to 1% of the average actual value for that state, then the estimate for that state is declared acceptable. We use the percentage of estimates considered acceptable (PEA) in our comparisons.

The third factor is numerical stability. The numerical instability measure (NIM) is defined as the number of trials in which singularities are found divided by the total number of trials. When calculating the Monte Carlo average for each simulation it is reasonable to exclude any trials in which numerical instability is observed.

The results of this particular experiment suggest that the EKF is the best filter. However, results from other experiments show that the performance of a filter is dependent on the object trajectory and other factors. We provide evidence of this claim in the the following sections.

4.3 Adding Optical Flow

In this section, we examine the effect of including optical flow measurements in the observations. Starting with initial estimates that have errors ranging from 20% to 40% and without using optical flow measurements, the EKF produces better estimates than the IEKF and ILFS. Once the optical flow measurements have been added, however, the ILFS emerges as the best among the three filters. As seen from Figure 1 and Table 2, the extra information provided by the optical flow measurements results in noticeable improvements. However, results from other experiments indicate that the proposed hybrid feature/flow-based approach does not work well when the initial state estimates are "crude" (e.g., when errors are between 50% and 100% or when no initial information is given). This finding suggests that a batch algorithm, such as the one outlined in [1], should be used to obtain reasonably reliable initial estimates. Without considering experiments where no initial information is given, the EKF has an overall NIM of 0.127, compared to 0.011 and 0.018 for the IEKF and ILFS, respectively. In other words, the EKF is numerically more unstable than the IEKF and ILFS when optical flow measurements are used.

4.4 Making Use of Process Noise

Although our model explicitly states that the object motion consists of constant translational and rotational velocities only, it is likely that we will encounter

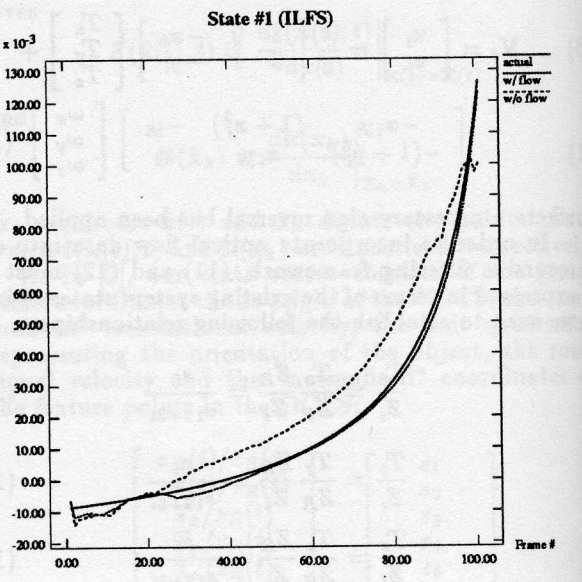

Figure 1: The effect of adding optical flow

Flow	x_R/z_R	y_R/z_R	T_x/z_R	T_y/z_R
No	1.3×10^{-4}	8.5×10^{-5}	1.0×10^{-8}	1.0×10^{-8}
Yes	4.1×10^{-6}	7.8×10^{-6}	1.0×10^{-8}	0

Flow	T_z/z_R	ω_x	ω_y	ω_z
No	3.2×10^{-6}	7.6×10^{-6}	3.9×10^{-6}	7.5×10^{-6}
Yes	1.7×10^{-6}	6.5×10^{-7}	2.4×10^{-6}	1.7×10^{-6}

Table 2: The effect of adding optical flow (MSE)

	EKF	IEKF	ILFS
BC	0	5	26
PEA	42.5%	99.6%	99.6%

Table 3: Filter comparison (optical flow)

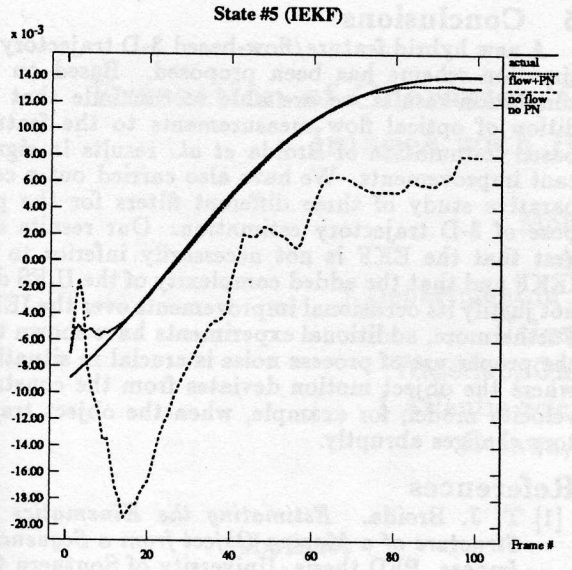

Figure 2: Translational acceleration

Flow/PN	x_R/z_R	y_R/z_R	T_x/z_R	T_y/z_R
No/No	5.6×10^{-3}	5.0×10^{-3}	1.2×10^{-4}	7.9×10^{-5}
No/Yes	8.1×10^{-5}	6.5×10^{-5}	1.0×10^{-5}	1.2×10^{-5}
Yes/Yes	2.0×10^{-6}	3.4×10^{-6}	1.0×10^{-7}	1.2×10^{-7}
Flow/PN	T_z/z_R	ω_x	ω_y	ω_z
No/No	7.3×10^{-5}	4.2×10^{-4}	1.3×10^{-4}	8.4×10^{-5}
No/Yes	3.4×10^{-5}	2.1×10^{-6}	1.6×10^{-6}	9.9×10^{-6}
Yes/Yes	1.4×10^{-7}	2.5×10^{-6}	2.6×10^{-7}	7.9×10^{-7}

Table 4: Translational acceleration (MSE)

	EKF	IEKF	ILFS
BC	0	24	14
PEA	100%	100%	100%

Table 5: Filter comparison (acceleration)

less-than-ideal situations in practice. The use of process noise (PN), in theory, can compensate for such modelling errors. Unfortunately, since no systematic procedure exists for determining the "proper amount" of process noise, its magnitude is selected based on a trial-and-error approach. For simplicity, the process noise is assumed to be constant over time.

Two sets of experiments are presented in this section. The first set consists of an object rotating at a velocity while changing its translational velocity at a constant rate. This clearly violates our previous assumption of constant translational velocity. In this case, both the IEKF and ILFS perform substantially better than the EKF, with the IEKF giving slightly better results than the ILFS. Simulation results also show that the addition of process noise alone is not sufficient. It is only when both process noise and optical flow measurements are utilized that significant improvements are achieved (see Table 4 and Figure 2). The acceleration used in this experiment is $\mathbf{a} = [0.01\ 0.01\ 0.01]$.

In the second set, the object motion deviates even more from the model as the object is allowed to change its trajectory *abruptly* after 50 frames. Here the emphasis is on the accuracy of the estimates made *after* the abrupt change; therefore, the definition of the mean-squared-error must be modified accordingly:

$$MSE(after) = \frac{1}{40}\sum_{i=61}^{100}(\hat{\mathbf{x}}_i - \mathbf{x}_i)^2 \qquad (23)$$

Due to fluctuations that are likely to occur immediately after the abrupt change, frames $51-60$ are excluded. Again, adding process noise alone does not provide satisfactory results; optical flow measurements must be used as well before any substantial improvements can be observed (see Table 6 and Figure 3).

Although the EKF does not perform as well as the IEKF or ILFS *before* the abrupt change, it yields the smallest MSE *after* the abrupt change. The motion parameters for frames $51-100$ are $\mathbf{T} = [0.1\ 0.12\ 0.025]$ and $\omega = [0.06\ 0.04\ 0]$.

4.5 Future Considerations

We have recently started studying the feasibility of removing the quaternion states. The key idea is that instead of propagating the rotation matrix using the unit quaternion, we use the rotational velocity vector as proposed by Young *et al.* [10]. Preliminary results suggest that the elimination of the quaternion states is feasible but further experimentation is necessary.

In our simulation, we avoid dealing with the problem of self-occlusion, or disappearance of feature points due to object's rotational motion. We simply treat the object as if it were transparent. One way to tackle this problem is to minimize the number of image frames in which we need to establish feature correspondences. Towards this goal, we have recently started investigating the use of two-view motion estimation [8]. The potential advantage of this technique is that there is no need to maintain feature correspondence over a large number of image frames.

We are also interested in discovering ways to obtain "pseudo-measurements" which allow us to observe certain states more directly. For example, direct observation of the inter-frame quaternion, translational direction and relative depth is possible using the two-view motion estimation algorithm proposed by Weng *et al.* [8]. More recently, Heeger and Jepson [4] proposed a method for computing motion and depth from optical flow data. The use of pseudo-measurements also leads to the elimination of the structure states and thus reduces the size of the state vector.

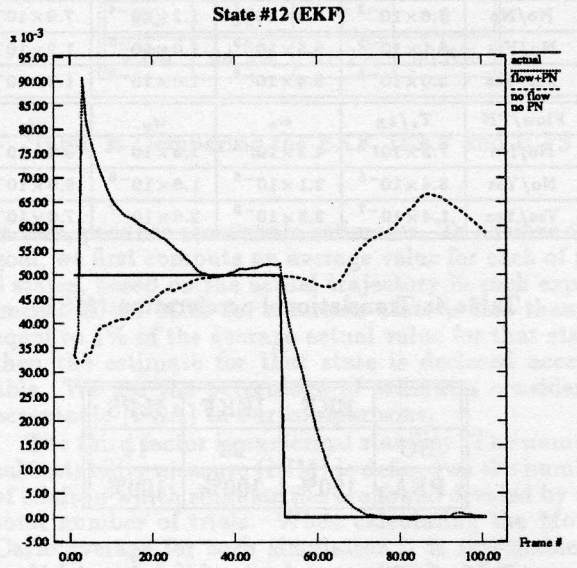

Figure 3: Abrupt change in object trajectory

Flow/PN	x_R/z_R	y_R/z_R	T_x/z_R	T_y/z_R
No/No	3.9×10^{-3}	2.1×10^{-2}	5.2×10^{-6}	1.2×10^{-6}
No/Yes	1.8×10^{-4}	5.3×10^{-4}	1.2×10^{-6}	1.8×10^{-6}
Yes/Yes	5.6×10^{-6}	3.5×10^{-6}	1.2×10^{-7}	1.0×10^{-7}
Flow/PN	T_z/z_R	ω_x	ω_y	ω_z
No/No	1.9×10^{-4}	4.2×10^{-4}	5.1×10^{-5}	3.5×10^{-3}
No/Yes	2.5×10^{-6}	7.6×10^{-6}	5.2×10^{-6}	3.9×10^{-5}
Yes/Yes	2.4×10^{-7}	9.4×10^{-7}	1.6×10^{-6}	9.4×10^{-6}

Table 6: Abrupt change in object trajectory (MSE)

	EKF	IEKF	ILFS
BC	30	0	0
PEA	99.2%	89.2%	89.2%

Table 7: Filter comparison (abrupt change)

5 Conclusions

A new hybrid feature/flow-based 3-D trajectory estimation scheme has been proposed. Based on our simulation results we are able to conclude that addition of optical flow measurements to the feature-based formulation of Broida *et al.* results in significant improvements. We have also carried out a comparative study of three different filters for the purpose of 3-D trajectory estimation. Our results suggest that the EKF is not necessarily inferior to the IEKF and that the added complexity of the ILFS does not justify its occasional improvements over the IEKF. Furthermore, additional experiments have shown that the proper use of process noise is crucial in situations where the object motion deviates from the constant-velocity model, for example, when the object trajectory changes abruptly.

References

[1] T. J. Broida. *Estimating the Kinematics and Structure of a Moving Object from a Sequence of Images.* PhD thesis, University of Southern California, 1987.

[2] T. J. Broida, S. Chandrashekhar, and R. Chellappa. Recursive 3-D Motion Estimation from a Monocular Image Sequence. *IEEE Trans. AES*, 26(4):639–656, 1990.

[3] A. R. Bruss and B. K. P. Horn. Passive Navigation. *CVGIP*, 21(1):3–20, 1983.

[4] D. J. Heeger and A. D. Jepson. Subspace Methods for Recovering Rigid Motion I:Algorithm and Implementation. *IJCV*, 7(2):95–117, 1992.

[5] S. Iu and K. Wohn. Estimation of General Rigid Body Motion From a Long Sequence of Images. In *Proc. IEEE 10th ICPR*, pages 217–219, 1990.

[6] A. H. Jazwinski. *Stochastic Processes and Filtering Theory.* Academic Press, 1970.

[7] O. Silvén and T. Repo. Solutions for Real-Time Visual Tracking. In *SPIE Vol. 1708 Applications of Artificial Intelligence X: Machine Vision and Robotics*, pages 184–195. International Society for Optical Engineering, 1992.

[8] J. Weng, T. S. Huang, and N. Ahuja. Motion and Structure from Two Perspective Views: Algorithms, Error Analysis, and Error Estimation. *IEEE Trans. PAMI*, 11(5):451–476, 1989.

[9] R. P. Wishner, J. A. Tabaczynski, and M. Athans. A Comparison of Three Non-Linear Filters. *Automatica*, 5:487–496, 1969.

[10] G. S. Young, R. Chellappa, and T. H. Wu. Monocular Motion Estimation Using a Long Sequence of Noisy Images. In *Proc. IEEE ICASSP '91*, pages 2437–2440, 1991.