

A Filter for Visual Tracking Based on a Stochastic Model for Driver Behaviour

Stephen J. Maybank, Anthony D. Worrall and Geoffrey D. Sullivan

Dept. of Computer Science, University of Reading, RG6 6AY, UK.
Email: (S.J.Maybank, A.D.Worrall, G.D.Sullivan)@reading.ac.uk.

Abstract. A driver controls a car by turning the steering wheel or by pressing on the accelerator or the brake. These actions are modelled by Gaussian processes, leading to a stochastic model for the motion of the car. The stochastic model is the basis of a new filter for tracking and predicting the motion of the car, using measurements obtained by fitting a rigid 3D model to a monocular sequence of video images. Experiments show that the filter easily outperforms traditional filters.

1 Introduction

The abilities to track and to predict car motion are important in any vision system for road traffic monitoring. Accurate tracking is required to simplify the measurement process. Tracking and prediction rely on filters which contain a model for the vehicle motion. Traditional filters such as the extended Kalman filter (EKF) [4] or the $\alpha - \beta$ filter [1] have very poor models for the car motion. For this reason they perform badly whenever the car carries out a complicated manoeuvre, for example a three point turn. This paper describes a new filter which contains a simple but accurate model for car motion. Experiments show that the filter easily outperforms the $\alpha - \beta$ filter.

The model is based on the behaviour of a driver who controls the motion of a car by turning the steering wheel, by pressing the accelerator or the brake, and by changing to a reverse gear, as appropriate. The velocity of the car depends on the steering angle θ and the signed magnitude v of the velocity. The position and orientation of the car are obtained by integrating the velocity over time.

The new filter uses a Gaussian approximation to the probability density function for the state of the car, conditional on the measurements. This density is propagated forward in time using the model for the motion of the car. Measurements of the position and orientation of the car are obtained by fitting the projection of a rigid 3D wire frame to individual video images taken by a single fixed camera; the measurements are incorporated into the approximating density using the standard Kalman filter update.

The treatment developed here is fundamentally different from that of other stochastic filters such as the Kalman filter, in which the state variables are assumed to have a linear interaction. In the new filter the covariances of the state variables are updated with errors of order only $O(t^3)$, where t is the time step. For this reason, we call the filter the Covariance-Updating (CU) filter.

In comparison with the $\alpha - \beta$ filter, the CU filter has a greatly improved response to changes of state which are not explicitly included in the model, for example changes involving higher order terms, which are only modelled implicitly as process noise. Thus, the CU filter is more robust when the vehicle switches from a constant forward acceleration to a turning or a stopping motion.

2 The Filter

The CU filter is based on a motion model which enforces the links between the driver behaviour and the car motion, and which also enforces the simple geometric constraints applicable to car motion.

2.1 Driver Behaviour and the Motion of the Car

It is assumed that the car driver can vary the steering angle θ , and the signed magnitude v of the velocity of the car. Changes in θ are achieved by turning the steering wheel and changes in v are achieved by pressing on the accelerator or the brake, or by using the reverse gear. In the absence of any further prior knowledge, the changes over time in θ , v are modelled by Gaussian stochastic processes Θ , V in \mathbb{R} defined by

$$(\Theta_t, V_t)^\top = (\Theta_0, V_0)^\top + \int_0^t (\dot{\Theta}_s, \dot{V}_s)^\top ds \quad (0 \leq t) \quad (1)$$

The random variables Θ_0 , V_0 in (1) are Gaussian. The rate of change of steering angle, $\dot{\Theta}$, is an Ornstein-Uhlenbeck process [2] which satisfies the Itô stochastic differential equation

$$d\dot{\Theta}_t = -\alpha\dot{\Theta}_t dt + \sigma dC_t \quad (0 \leq t) \quad (2)$$

where $\alpha > 0$, $\sigma > 0$ are constants, and C is a Brownian motion in \mathbb{R} such that $C_0 = 0$. The random variable $\dot{\Theta}_0$ is Gaussian and independent of C . The process \dot{V} in (1) is given by

$$\dot{V}_t = \dot{V}_0 + qB_t \quad (0 \leq t) \quad (3)$$

where $q > 0$ is constant and B is a Brownian motion in \mathbb{R} independent of C and $B_0 = 0$. The random variable \dot{V}_0 is Gaussian and independent of B , C .

The parameters σ , α , q have the following effects: if σ , q are large, then the variances of $\dot{\Theta}_t$, \dot{V}_t tend to be large. The filter becomes more responsive to recent measurements and rapidly discards past measurements. If α is large then the variance of $\dot{\Theta}_t$ tends to be small, and the estimates of Θ_t , Θ_t become less responsive to all the measurements. At the same time the trajectory of the car, as generated by (5), shows a reduced curvature.

It remains to link the driver behaviour to the motion of the car. It is assumed that the car moves in the ground plane, and that the motion is completely described by a time dependent state vector $(x, y, \theta, v, \dot{\theta}, \dot{v})^\top$, where $(x, y)^\top$ is the

position of the car in the ground plane, θ is the orientation of the car, v is the signed magnitude of the velocity of the car, and $\dot{\theta}$, \dot{v} are the time derivatives of θ , v respectively.

The model for the time evolution of the state vector is a stochastic process

$$M = (X, Y, \theta, V, \dot{\theta}, \dot{V})^\top \quad (4)$$

in \mathbb{R}^6 , where $(X, Y)^\top$ models the position of the vehicle on the ground plane, and $\theta, V, \dot{\theta}, \dot{V}$ are as defined by (1), (2) and (3). The process $(X, Y)^\top$ is given in terms of $(\theta, V)^\top$ by

$$(X_t, Y_t)^\top = (X_0, Y_0)^\top + \int_0^t V_s (\cos(\theta_s), \sin(\theta_s))^\top ds \quad (0 \leq t) \quad (5)$$

Equation (5) enforces a fundamental geometric constraint: relative to the car body, the direction of the velocity is fixed up to sign (cars cannot move sideways). In this case the direction of the velocity $(\dot{X}_t, \dot{Y}_t)^\top$ fixes the steering angle θ_t through the equations $\dot{X}_t = V_t \cos(\theta_t)$, $\dot{Y}_t = V_t \sin(\theta_t)$ for $0 \leq t$.

In practice, v and $\dot{\theta}$ are correlated for v small: it is difficult to steer rapidly a slowly moving car. The performance of the CU filter was improved by making α a function of the expected tangential velocity \hat{v} , $\alpha = 0.01 + 2.0 \exp(-\hat{v}^2/2)$. At high \hat{v} , α is small, to ensure that the filter can track a car successfully while it is turning. At low values of \hat{v} , α is large, thus dragging $\dot{\theta}$ towards zero. Very large values of α are avoided, to prevent the filter locking on to an erroneous estimate of θ .

2.2 Propagation of the Density and the Measurement Update

The task of the filter is to estimate the probability density function for the state M_t of the car at a time t , given the measurements obtained at times prior to and including t . In this application the measurements are of x, y, θ . It is assumed that the measurement errors are Gaussian with zero expected value and a known covariance matrix.

The strategy in the filter is to approximate the true conditional density with a Gaussian density. The approximation is reliable provided the pose of the car is well localised by the measurements. The filter propagates the approximating density forward in time, incorporating each measurement as it arises. It is convenient to discuss propagation from time zero, to reduce the notation. Let p_0 be the Gaussian approximation to the density of M_0 . Then it is required to find a Gaussian approximation p_t to the density of M_t , $t \geq 0$, assuming that no measurements are taken in the interval $[0, t)$. If t is small, then the expectation and covariance of p_t are calculated correct to $O(t^3)$ using the formulae given in the next section. If t is large, then $[0, t)$ is divided into small subintervals $0 = t_1 < \dots < t_n = t$, and the p_{t_i} are calculated in turn until p_t is reached.

A model of the vehicle is instantiated into the image in the pose predicted by the current state of the filter, as a wire-frame with hidden lines removed. Evidence for directed edges in the image that fall near to the visible "wires"

is pooled to provide a scalar value associated with the pose (see [6] and [8] for further details). This defines a function on the space of vehicle poses. A search is carried out to find a local maximum of the function under small perturbations of the pose from the initial pose, using the simplex search technique [5]. The recovered location $(x, y)^\top$ and orientation θ of the vehicle form the measurement vector z .

Let a measurement z_t be obtained at time $t \geq 0$. The measurement update is the standard one used in the Kalman filter [3]. Let the estimated density p_t prior to the measurement have expectation m_t and covariance C_t . Let $w \mapsto Hw$, $w \in \mathbb{R}^6$ be the measurement function, and let R be the covariance matrix for the measurement errors, which are assumed to be Gaussian with expectation zero. The 3×6 matrix H is defined by $H_{ii} = 1$, $1 \leq i \leq 3$ and $H_{ij} = 0$, $i \neq j$. The expected value n_t and the covariance matrix D_t of the estimated density of M_t conditional on z_t are fixed by

$$(w - n_t)^\top D_t^{-1} (w - n_t) \equiv (w - m_t)^\top C_t^{-1} (w - m_t) + (z_t - Hw)^\top R^{-1} (z_t - Hw) + c_t$$

where c_t is independent of w .

3 Propagation of Expectations and Covariances

Let p_0 be the Gaussian approximation to the density of M_0 . A Gaussian approximation p_t to the density of M_t is obtained under the assumption that $t > 0$ is small. The expectation and covariance of p_t are correct to $O(t^3)$, provided the density of M_0 is exactly p_0 .

3.1 Expectation and Covariance of $(\Theta_t, V_t, \dot{\Theta}_t, \dot{V}_t)^\top$

The expected values and the covariances of $\Theta_t, V_t, \dot{\Theta}_t, \dot{V}_t$ are easy to evaluate, in part because $\dot{\Theta} - \dot{\Theta}_0$ and $\dot{V} - \dot{V}_0$ are independent Gaussian processes. It follows from (2) and (3) that

$$\begin{aligned} E(\dot{\Theta}_t, \dot{V}_t) &= E(\exp(-\alpha t) \dot{\Theta}_0, \dot{V}_0) \\ \text{Cov}(\dot{\Theta}_t, \dot{\Theta}_t) &= \exp(-2\alpha t) \text{Cov}(\dot{\Theta}_0, \dot{\Theta}_0) + \alpha^{-1} \sigma^2 \exp(-\alpha t) \sinh(\alpha t) \\ \text{Cov}(\dot{V}_t, \dot{V}_t) &= \text{Cov}(\dot{V}_0, \dot{V}_0) + q^2 t \\ \text{Cov}(\dot{\Theta}_t, \dot{V}_t) &= \exp(-\alpha t) \text{Cov}(\dot{\Theta}_0, \dot{V}_0) \end{aligned} \quad (6)$$

Let $r = t - \alpha t^2/2$. It follows from (1) and (6) that

$$\begin{aligned} E(\Theta_t) &= E(\Theta_0) + E(\dot{\Theta}_0)r + O(t^3) \\ E(V_t) &= E(V_0) + E(\dot{V}_0)t \\ \text{Cov}(\Theta_t, \Theta_t) &= \text{Cov}(\Theta_0, \Theta_0) + 2\text{Cov}(\Theta_0, \dot{\Theta}_0)r + \text{Cov}(\dot{\Theta}_0, \dot{\Theta}_0)t^2 + O(t^3) \\ \text{Cov}(V_t, V_t) &= \text{Cov}(V_0, V_0) + 2\text{Cov}(V_0, \dot{V}_0)t + \text{Cov}(\dot{V}_0, \dot{V}_0)t^2 + O(t^3) \\ \text{Cov}(\Theta_t, V_t) &= \text{Cov}(\Theta_0, V_0) + \text{Cov}(\dot{\Theta}_0, V_0)r + \text{Cov}(\Theta_0, \dot{V}_0)t + \text{Cov}(\dot{\Theta}_0, \dot{V}_0)t^2 \end{aligned}$$

The cross covariances between $(\Theta_t, V_t)^\top$ and $(\dot{\Theta}_t, \dot{V}_t)^\top$ are

$$\begin{aligned} \text{Cov}(\dot{\Theta}_t, \Theta_t) &= \exp(-\alpha t)(\text{Cov}(\dot{\Theta}_0, \Theta_0) + \text{Cov}(\dot{\Theta}_0, \dot{\Theta}_0)r + \sigma^2 t^2/2) \\ \text{Cov}(\dot{V}_t, \Theta_t) &= \text{Cov}(\dot{V}_0, \Theta_0) + \text{Cov}(\dot{V}_0, \dot{\Theta}_0)r \\ \text{Cov}(\dot{\Theta}_t, V_t) &= \exp(-\alpha t)(\text{Cov}(\dot{\Theta}_0, V_0) + \text{Cov}(\dot{\Theta}_0, \dot{V}_0)t) \\ \text{Cov}(\dot{V}_t, V_t) &= \text{Cov}(\dot{V}_0, V_0) + \text{Cov}(\dot{V}_0, \dot{V}_0)t + 2^{-1}q^2 t^2 \end{aligned}$$

3.2 Expectation of $(X_t, Y_t)^\top$

The covariances and expectations involving X_t, Y_t are more difficult to evaluate than those involving only $\Theta_t, V_t, \dot{\Theta}_t, \dot{V}_t$. The calculations are reduced by combining X and Y to make a complex valued stochastic process, $Z = X + iY$, where $i^2 = -1$. It follows from (5) that to an error $O(t^3)$,

$$Z_t = Z_0 + \exp(i\Theta_0) \left(V_0 t + 2^{-1} \dot{V}_0 t^2 \right) + 2^{-1} i \exp(i\Theta_0) V_0 \dot{\Theta}_0 t^2 \quad (0 \leq t) \quad (7)$$

Let $c = E(\exp(i\Theta_0))$. It follows from (7) that

$$\begin{aligned} \frac{E(Z_t)}{c} &= \frac{E(Z_0)}{c} + t[i\text{Cov}(\Theta_0, V_0) + E(V_0)] + \frac{t^2}{2}[i\text{Cov}(\Theta_0, \dot{V}_0) + E(\dot{V}_0)] \\ &\quad + \frac{it^2}{2}[\text{Cov}(V_0, \dot{\Theta}_0) - \text{Cov}(\Theta_0, V_0)\text{Cov}(\Theta_0, \dot{\Theta}_0) + iE(V_0)\text{Cov}(\Theta_0, \dot{\Theta}_0)] \\ &\quad - \frac{t^2}{2}[E(\dot{\Theta}_0)\text{Cov}(\Theta_0, V_0) - iE(V_0)E(\dot{\Theta}_0)] + O(t^3) \end{aligned} \quad (8)$$

3.3 Covariances Involving X_t or Y_t

To simplify the notation, let $W \mapsto g(W)$ be the function defined on Gaussian random variables by

$$g(W) = \text{Cov}(Z_t, W) \quad (0 \leq t) \quad (9)$$

It follows from (7) and (9) that

$$\begin{aligned} \frac{g(W)}{c} &= \frac{\text{Cov}(Z_0, W)}{c} + t[\text{Cov}(V_0, W) - \text{Cov}(\Theta_0, V_0)\text{Cov}(\Theta_0, W)] \\ &\quad + itE(V_0)\text{Cov}(\Theta_0, W) + 2t^2[\text{Cov}(\dot{V}_0, W) - \text{Cov}(\Theta_0, \dot{V}_0)\text{Cov}(\Theta_0, W)] \\ &\quad + 2t^2[iE(\dot{V}_0)\text{Cov}(\Theta_0, W) - \text{Cov}(\dot{\Theta}_0, W)\text{Cov}(\Theta_0, V_0)] \\ &\quad - 2t^2[\text{Cov}(V_0, \dot{\Theta}_0)\text{Cov}(\Theta_0, W) + \text{Cov}(W, V_0)\text{Cov}(\Theta_0, \dot{\Theta}_0)] \\ &\quad + 2t^2[\text{Cov}(\Theta_0, W)\text{Cov}(\Theta_0, \dot{\Theta}_0)\text{Cov}(\Theta_0, V_0) + iE(V_0)\text{Cov}(\dot{\Theta}_0, W)] \\ &\quad - 2it^2[E(V_0)\text{Cov}(\Theta_0, \dot{\Theta}_0)\text{Cov}(\Theta_0, W) + E(\dot{\Theta}_0)\text{Cov}(\Theta_0, V_0)\text{Cov}(\Theta_0, W)] \\ &\quad + 2t^2[iE(\dot{\Theta}_0)\text{Cov}(V_0, W) - E(V_0)E(\dot{\Theta}_0)\text{Cov}(\Theta_0, W)] + O(t^3) \end{aligned}$$

The cross covariances of Z_t with $\Theta_t, V_t, \dot{\Theta}_t, \dot{V}_t$ are (to $O(t^3)$)

$$\begin{aligned} \text{Cov}(Z_t, \Theta_t) &= g(\Theta_0) + g(\dot{\Theta}_0)t & \text{Cov}(Z_t, \dot{\Theta}_t) &= g(\dot{\Theta}_0) \\ \text{Cov}(Z_t, V_t) &= g(V_0) + g(\dot{V}_0)t & \text{Cov}(Z_t, \dot{V}_t) &= g(\dot{V}_0) \end{aligned} \quad (10)$$

It remains to find the covariance of $(X_t, Y_t)^\top$. To simplify the notation, let the functions $W \mapsto f_1(W)$ and $W \mapsto f_2(W)$, $\Theta \mapsto h(\Theta)$ of the Gaussian random variables W, Θ be defined by

$$\begin{aligned} f_1(W) &= \text{Cov}(Z_0, \exp(i\Theta_0)W) \\ &\equiv c[\text{Cov}(W, Z_0) - \text{Cov}(W, \Theta_0)\text{Cov}(Z_0, \Theta_0) + iE(W)\text{Cov}(Z_0, \Theta_0)] \\ f_2(W) &= \text{Cov}(Z_0, \exp(-i\Theta_0)W) \\ &\equiv \bar{c}[\text{Cov}(W, Z_0) - \text{Cov}(W, \Theta_0)\text{Cov}(Z_0, \Theta_0) - iE(W)\text{Cov}(Z_0, \Theta_0)] \\ h(\Theta) &= \text{Cov}(Z_0, \exp(i\Theta)V_0\dot{\Theta}_0) \end{aligned} \quad (11)$$

Let $c_\theta = E(\exp(i\theta))$. It follows from the last equation of (11) that

$$\begin{aligned} c_\theta^{-1}h(\theta) &= i[\text{Cov}(Z_0, V_0)\text{Cov}(\theta, \dot{\Theta}_0) + \text{Cov}(Z_0, \dot{\Theta}_0)\text{Cov}(\theta, V_0)] \\ &\quad + i[\text{Cov}(Z_0, \theta)\text{Cov}(V_0, \dot{\Theta}_0) - \text{Cov}(Z_0, \theta)\text{Cov}(\theta, \dot{\Theta}_0)\text{Cov}(\theta, V_0)] \\ &\quad + E(V_0)[\text{Cov}(Z_0, \dot{\Theta}_0) - \text{Cov}(Z_0, \theta)\text{Cov}(\theta, \dot{\Theta}_0)] \\ &\quad + E(\dot{\Theta}_0)[\text{Cov}(Z_0, V_0) - \text{Cov}(Z_0, \theta)\text{Cov}(\theta, V_0)] \\ &\quad + iE(V_0)E(\dot{\Theta}_0)\text{Cov}(Z_0, \theta) \end{aligned} \quad (12)$$

Let A, B be the random variables defined by

$$A = \exp(i\Theta_0)V_0 \quad B = \exp(i\Theta_0)(\dot{V}_0 + iV_0\dot{\Theta}_0) \quad (13)$$

It follows from (7) and (13) that

$$Z_t = Z_0 + At + Bt^2/2 + O(t^3) \quad (0 \leq t) \quad (14)$$

thus (to order $O(t^3)$)

$$\begin{aligned} \text{Cov}(Z_t, Z_t) &= \text{Cov}(Z_0, Z_0) + 2t\text{Cov}(Z_0, A) + t^2[\text{Cov}(Z_0, B) + \text{Cov}(A, A)] \\ &= \text{Cov}(Z_0, Z_0) + 2tf_1(V_0) + t^2[f_1(\dot{V}_0) + ih(\Theta_0) + \text{Cov}(A, A)] \\ \text{Cov}(\bar{Z}_t, Z_t) &= \text{Cov}(\bar{Z}_0, Z_0) + t(\text{Cov}(\bar{Z}_0, A) + \text{Cov}(\bar{A}, Z_0)) + t^2\text{Cov}(\bar{A}, A) \\ &\quad + 2^{-1}t^2[\text{Cov}(\bar{Z}_0, B) + \text{Cov}(\bar{B}, Z_0)] \\ &= \text{Cov}(\bar{Z}_0, Z_0) + 2t\text{Re}(f_2(V_0)) + t^2\text{Re}(f_2(\dot{V}_0) - ih(-\Theta_0)) \\ &\quad + t^2\text{Cov}(\bar{A}, A) \end{aligned} \quad (15)$$

A short calculation yields

$$\begin{aligned} \text{Cov}(A, A) &= c^3\bar{c}[\text{Cov}(V_0, V_0) - 4\text{Cov}(\Theta_0, V_0)^2 + 4iE(V_0)\text{Cov}(\Theta_0, V_0) + E(V_0)^2] \\ &\quad + c^2[\text{Cov}(\Theta_0, V_0)^2 - 2iE(V_0)\text{Cov}(\Theta_0, V_0) - E(V_0)^2] \\ \text{Cov}(\bar{A}, A) &= \text{Cov}(V_0, V_0) + E(V_0)^2 - c\bar{c}(\text{Cov}(\Theta_0, V_0)^2 + E(V_0)^2) \end{aligned} \quad (16)$$

The covariances $\text{Cov}(Z_t, Z_t)$, $\text{Cov}(\bar{Z}_t, Z_t)$ can be evaluated to an accuracy of $O(t^3)$ using (11), (12), (15) and (16). The covariance of $(X_t, Y_t)^\top$ is given by

$$\begin{aligned} \text{Cov}(X_t, X_t) &= 2^{-1}[\text{Cov}(\bar{Z}_t, Z_t) + \text{Re}(\text{Cov}(Z_t, Z_t))] \\ \text{Cov}(X_t, Y_t) &= 2^{-1}\text{Im}(\text{Cov}(Z_t, Z_t)) \\ \text{Cov}(Y_t, Y_t) &= 2^{-1}[\text{Cov}(\bar{Z}_t, Z_t) - \text{Re}(\text{Cov}(Z_t, Z_t))] \end{aligned}$$

3.4 Comparison with the Extended Kalman Filter

The extended Kalman filter (EKF) is a popular generalisation of the Kalman filter to nonlinear systems [4]. In spite of its widespread use, the EKF often gives an astonishingly bad approximation to the true density of the system state conditional on the measurements. The EKF differs radically from the CU filter, for example the latter is accurate to within an error of $O(t^3)$. In the EKF the estimate of $E(M_t)$ is independent of $\text{Cov}(M_0, M_0)$. However, it is clear from (5) that X_t, Y_t depend on products of functions of $\Theta_s, V_s, 0 \leq s \leq t$. It follows that an accurate approximation to $E(M_t)$ includes terms depending on the covariances of $\Theta_s, V_s, 0 \leq s \leq t$. In fact, the covariances are needed even if the approximation is correct only to $O(t^2)$.

In more detail, let M_0 have a Gaussian density, and let $Z_t = X_t + iY_t$. It follows from (7) that the estimate \hat{Z}_t of $E(Z_t)$ produced by the EKF is

$$\hat{Z}_t = E(Z_0) + \exp(iE(\Theta_0))E(V_0)t + O(t^2)$$

As an approximation to $E(Z_t)$, \hat{Z}_t is incorrect to first order in t , because

$$E(Z_t) = E(Z_0) + E(\exp(i\Theta_0))(i\text{Cov}(\Theta_0, V_0) + E(\dot{V}_0))t + O(t^2)$$

4 Vehicle Tracking

The CU filter for driving behaviour has been implemented in Mathematica [7], and a C-code procedure was obtained, using the `Splice` command in Mathematica. The C code was incorporated, after trivial changes of syntax, into the pop-11 code used to control the model-based tracking system. The performance of the CU filter was then assessed in comparison with a simple $\alpha - \beta$ filter, by applying both filters to a video sequence of a car performing a three-point turn in a cluttered scene (see Fig. 1). In the past this sequence has proved unusually difficult to track, largely because of the complexity of the manoeuvre. In the absence of an accurate dynamical model, the measurement process is easily distracted by the background clutter.

Figure 1 (left) shows the trace on the ground-plane of the estimated trajectory, using the CU filter (grey line) updated by measurements (dots) taken at 25 Hz. The three outlying measurements near frame 120 were gated automatically using the Mahalanobis distance with a threshold of 4 standard deviations. Examples of the fit of the model to the image are also given in Fig. 1. The performance seems to be very good; the filter accurately followed the measurements, with significant smoothing and minimal overshoot at the cusps of the track. These filter states are taken as the "ground-truth" in subsequent experiments.

Figure 2 (top) shows the effects of using the CU filter to track the vehicle with increasing time gaps between the measurements for each run, using measurement intervals of 1, 2, 4 and 8 frames (at 25Hz). The propagation of the density was at 25Hz as described in Sect. 2.2. In each run the differences between the filter prediction and the ground truth are shown for the state variables x, y and

θ . It can be seen that the predictions usually stay very close to the ground-truth (ordinate = 0.0 in each graph). Some errors are seen up to frame 180, as the car turns sharply and comes to a halt, as well as between frames 300 and 400, where the car was viewed directly rear-on, and the measurement was more than usually unstable (see Fig. 1). The prediction errors became worse as the intervals between measurements increased. None the less, the predictions were always sufficiently accurate for the measurement process to recover a good solution, and the tracking succeeded.

Figure 2 (bottom) illustrates the performance of an α - β tracker, used to filter $(x, y, \theta, v, \dot{\theta})^\top$ with $\alpha = 0.5$ and $\beta = 0.1$. Performance is much poorer. Somewhat paradoxically, the best performance was obtained with measurements every second frame (dotted); the car was tracked successfully, though with noticeably more instability than with the CU filter. With measurements taken every frame (solid line), tracking failed towards the end (note the consistent errors in x and y beyond frame 380). With measurements every 4th frame (grey), tracking failed spectacularly; near frame 100 the recovered pose spun through 90° , and thereafter the car was tracked with the model at the wrong orientation.

4.1 Discussion

This paper has demonstrated a new method for building a dynamic filter to track vehicles in road scenes. The filter explicitly models driver behaviour by stochastic processes and it includes a realistic updating of the state covariances. When applied to pose data obtained from a model based vision system, initial results indicate that the covariance updating (CU) filter has significantly better performance than a simple linear filter such as the $\alpha - \beta$ filter.

The CU filter is fundamentally different from the usual stochastic filters such as the extended Kalman filter. The (assumed) dynamic model for the system is represented more accurately, because unlike the Kalman treatment, the evolution of the state variables is not assumed to be linear. The covariances which describe the nonlinear interactions between state variables are handled correctly up to and including terms of order $O(t^2)$, where t is a small time step. This greatly improves the performance of the filter when the evolution of the system is strongly nonlinear or when there is a short run of unreliable measurements.

References

1. Bar-Shalom, Y. and Fortmann, T.E.: Tracking and Data Association. Mathematics in Science and Engineering Series **179** (1988). Academic Press: San Diego, CA.
2. Kloeden, P.E., Platen, E.: Numerical Solution of Stochastic Differential Equations. Applications of Mathematics Series: stochastic modelling and applied probability **23** (1992). Springer-Verlag: Berlin.
3. Maybeck, P.S.: Stochastic Models, Estimation, and Control, **1**. Mathematics in Science and Engineering Series **141-1** (1979). Academic Press: San Diego, CA.
4. Maybeck, P.S.: Stochastic Models, Estimation, and Control **3**. Mathematics in Science and Engineering Series **141-3** (1982). Academic Press: San Diego, CA.

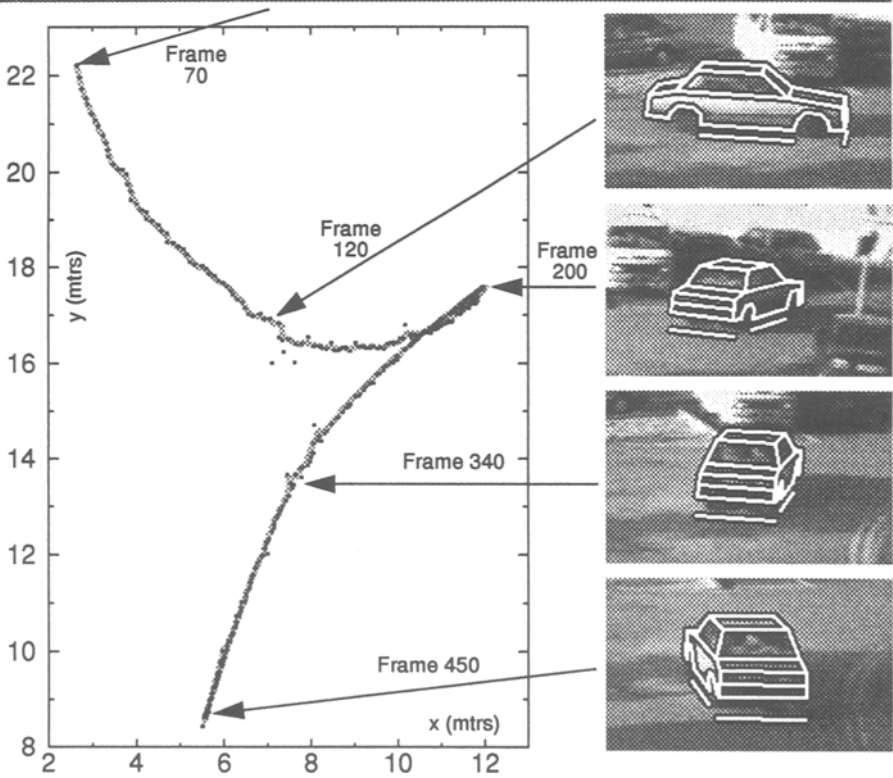


Figure 1 Top & right: Fragments of frames showing the car in ground-truth poses. Left: Trace of (x,y) ground-truth (grey line) and measured (dots) positions on the ground during 3-point turn manoeuvre.

5. Press, W.H.: Numerical Recipes. (1986) CUP: Cambridge, UK.
6. Sullivan, G.D.: Visual interpretation of known objects in constrained scenes. Phil. Trans. R. Soc. London, Series B 337 (1992) 361-370.

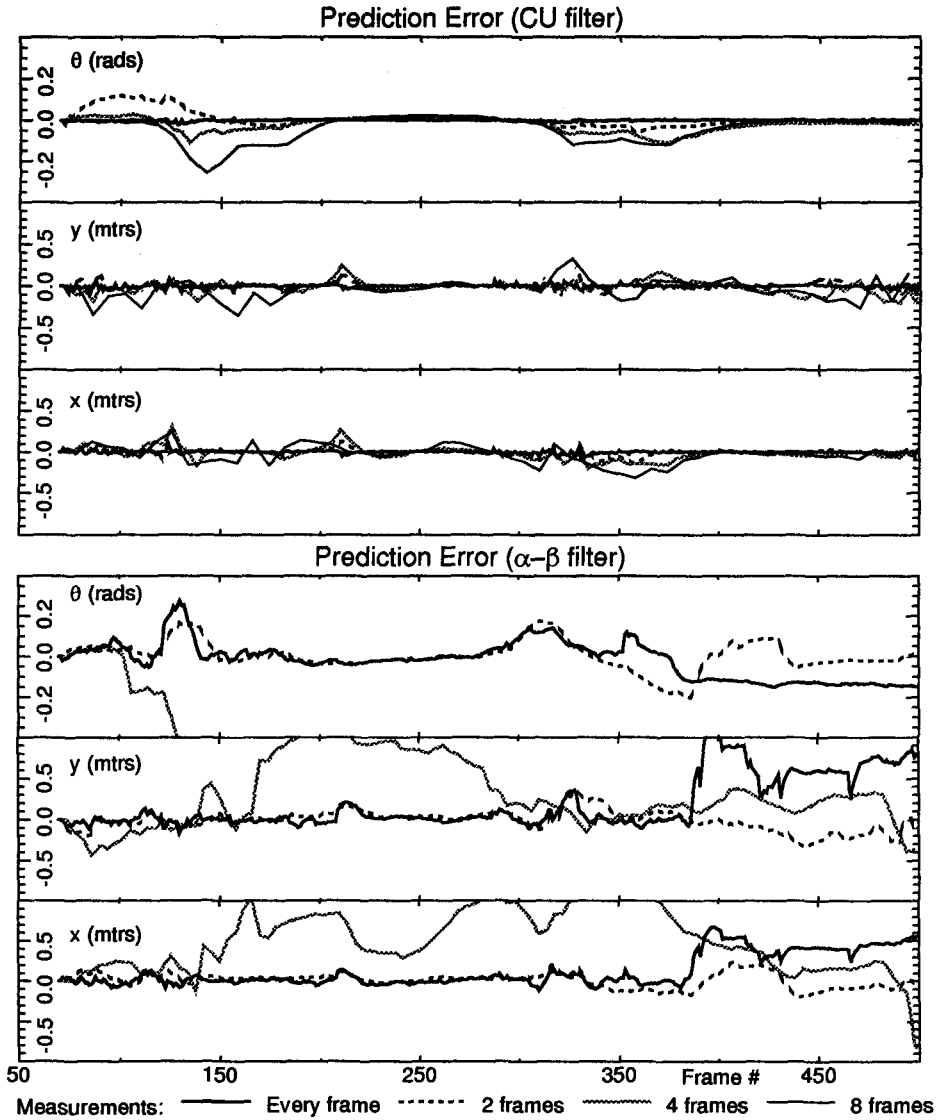


Figure 2: Prediction errors for pose parameters θ , x and y (w.r.t the ground-truth) as a function of frame number, using different measurement intervals.
 Top: The CU filter, for intervals of 1, 2, 4 and 8 frames (@ 25Hz).
 Bottom: The α - β filter, for intervals of 1, 2 and 4 frames (@ 25Hz).

7. Wolfram, S.: Mathematica, a system for doing mathematics by computer (1991) 2nd edition. Addison Wesley: Redwood City, CA.
8. Worrall, A.D., Ferryman, J.M., Sullivan, G.D. and Baker, K.D.: Pose and structure estimation using active models. Editor Pycocok, D.: British Machine Vision Conference 1995 1 (1995) 137-146. BMVA.