Tracking in a Hough Space with the Extended Kalman Filter

Steven Mills, Tony Pridmore, and Mark Hills
School of Computer Science and Information Technology
University of Nottingham
Nottingham, NG7 2RD, UK
{smx,tpp,mch}@cs.nott.ac.uk

Abstract

A combined tracking method using the Kalman filter and Hough transform is presented. An extended Kalman filter is used to model the parameters and motion of a set of lines detected in a Hough space The integration of these two techniques gives a number of advantages. The use of a Hough transform provides resilience to noise and partial occlusion, and the Kalman filter's ability to predict future states is used to reduce the computational load of line detection. Analysis of the tracker from synthetic data shows that it is robust to noise, occlusion, and deviations from the constant motion model underlying the Kalman filter. Tracking results from video sequences illustrate its applicability to real-world domains.

1 Introduction

In this paper we investigate the combined application of two established techniques in computer vision – the Hough transform [6] and the Kalman filter [8]. Although both of these techniques have received considerable attention and have been applied to a wide variety of application areas, there is little literature examining their combined application.

The Hough transform finds image features by collecting global evidence in a parameter space. In the most common application of the Hough transform the features are lines, and their parameters are their polar coordinates, ρ and θ [2]. Hough transforms can also be used to find circles, ellipses, and other shapes that can be represented by a small set of parameters. The Hough transform is, however, computationally expensive. It computes an accumulator array that holds the evidence for a feature for each set of parameters, if there are k parameters to be estimated, then this array is k-dimensional, and computing the array values from an image is $O(n^{k+1})$.

The Hough transform finds features in a single image, but in many applications we wish to track a feature, or set of features, over a series of images. Kalman filtering has emerged as one of the dominant techniques in this area. A Kalman filter is a recursively computed estimate of an object's state made on the basis of a series of measurements. In the case of tracking, the state of an object or image feature is its location and motion parameters. The Kalman filter also gives the covariance of this state estimate, making it possible to reason about errors and uncertainty in a principled manner.

The combination of the Hough transform and Kalman filter offers several advantages. The Hough transform is robust to image noise and partial occlusion, and so can provide measurements to the Kalman filter even when these are present in the image. Occlusion is of particular concern in image sequence analysis, since features can become occluded or disoccluded over time. The predictions and covariance estimates from the Kalman filter can be used to focus the search for image features to a small part of the accumulator array. This gives dramatic improvements in the efficiency of the Hough transform computations.

In Section 2 we discuss tracking in Hough spaces, and review previous approaches. The Kalman filter formulation for tracking lines in a Hough space is developed in Section 3. The use of a Kalman filter allows the Hough transform to be computed more efficiently, as discussed in Section 4. This section also presents results of experiments to evaluate the performance of the integrated technique. Results on real-world sequences are given in Section 5, followed by concluding remarks in Section 6.

2 Tracking in Hough Spaces

Previous work on tracking with Hough spaces have used variants of the Hough transform to compute motion parameters. The velocity Hough transform [9], for example, computes a Hough transform with additional dimensions for the motion of the features. This approach analyses the entire sequence and uses a high-dimensional Hough transform, so has a high computational load. Another variant of the Hough transform, the combinatorial Hough transform, has been applied to optical flow computation [10]. In this case brightness constancy assumptions [5] are used to constrain the flow, and the Hough transform is used to estimate the most likely motion parameters.

In this work we do not use the Hough transform as a tool for tracking, but aim to track features through their parameters represented by the Hough space. Hills *et al.* [4] track pairs of parallel lines through a Hough space by viewing the Hough accumulator as an approximation to the probabilistic Hough transform [11], and applying a joint probability function to represent the likelihood of a pair of points in the Hough space representing two parallel lines. This probability density function is then maximised to track linked features between two frames. This approach successfully tracks pairs of lines and rectangular objects, but has two main limitations. Firstly it is assumed that the lines are parallel (or at least that their relative angle is known), and secondly tracking is performed from frame to frame without any sense of history.

To overcome these limitations we propose a Kalman filter based technique that tracks peaks in the Hough space. The peak locations are the measurements which drive the Kalman filter, and the motion parameters of a set of lines are estimated. These lines can have any orientation, but are assumed to have the same motion model. Lines are represented by their polar representation, (ρ,θ), and are assumed to be rotating around some point (x,y) with rotational velocity ω. The lines and the centre of rotation are assumed to be moving with translational velocity (u,v). All of the lines share the same parameters, so are tracked as a group rather than individually.

The use of the Kalman filter to track in Hough space has several advantages. It allows a long term model to be built up, which can be used for prediction. Tracking from frame to frame does not provide such a model. Secondly it allows for more efficient computation of the Hough transform. Since the Kalman filter provides an estimate as to the next location of a line, along with the covariance of this estimate, it can be used to restrict the search for peaks in the Hough accumulator and, therefore, the computation of the accumulator. Finally the Kalman filter provides the ability to model specific situations with appropriate motion models, and to account for uncertainty in the model and observations.

The Kalman filter and Hough transform have previously been used together, but have not been as closely integrated. Behrens *et al.* [1] detect generalised cylinders, such as arteries in three-dimensional medical imagery, by tracking ellipses detected with a Hough transform. A three-dimensional image is seen as a sequence of two-dimensional slices, and cylindrical structures are detected as ellipses in each slice. These ellipses are tracked from slice to slice using the Kalman filter. A randomised Hough transform is used, and the predictions of the Hough transform are used to sample the data in a non-uniform manner.

The structures detected by Behrens *et al.* are quite specific in that they are described by a single structure having five parameters – the *x*- and *y*-coordinates of the centre of the ellipse, its axes, and the angle of rotation. This means that the complexity of the structures being tracked are translated in to computational complexity, as a five-dimensional Hough accumulator is used. In our method complex structures in the image are tracked as sets of simple features (lines in our examples) with a common motion model.

Foresti [3] also uses the Hough transform and Kalman filter together. Line segments detected in moving regions of images are tracked in order to describe the objects within those regions. These features are detected with a Hough transform, and their motion is used to provide measurements for an extended Kalman filter. This Kalman filter does not, however, track the lines in the Hough space. Instead, lines in one frame are matched to the line that best matches their parameters (including end points and length as well as orientation and location) in the next. Some limited feedback from the Kalman filtering to the line tracking is employed in cases of occlusion, but the goal of the Kalman filter is to estimate the depth of the lines on the basis of their motion in the image plane.

3 A Kalman Filter for Tracking in a Hough Space

The information provided to the Kalman filter is the locations of a set of n lines in each frame. These lines are assumed to be moving with constant translational and rotational velocity in the image plane, so the state vector for the Kalman filter consists of the coordinates, ρ_i, θ_i for $1 \leq i \leq n$, of each line and their motion parameters. These parameters are the centre of rotation, (x,y), the rotational velocity, ω, and the translational velocity (u,v). The states at subsequent times are related by

$$s_{t+1} = \begin{bmatrix} x_{t+1} \\ y_{t+1} \\ u_{t+1} \\ v_{t+1} \\ \omega_{t+1} \\ \rho_{1,t+1} \\ \theta_{1,t+1} \\ \vdots \\ \rho_{n,t+1} \\ \theta_{n,t+1} \end{bmatrix} = \begin{bmatrix} x_t + u_t \\ y_t + v_t \\ u_t \\ v_t \\ \omega_t \\ r(1,t) \\ \theta_{1,t} + \omega_t \\ \vdots \\ r(n,t) \\ \theta_{n,t} + \omega_t \end{bmatrix} = f(s_t), \tag{1}$$

where $r(i,t)$ is given by

$$\rho_{i,t} - x_t \cos(\theta_{i,t}) - y_t \sin(\theta_{i,t}) + (x_t + u_t)\cos(\theta_{i,t} + \omega_t) + (y_t + v_t)\sin(\theta_{i,t} + \omega_t). \quad (2)$$

At each time we measure the parameters of the lines with a a Hough transform giving

$$m_t = [\rho_{1,t}, \theta_{1,t}, \dots, \rho_{n,t}, \theta_{n,t}]^T = h(s_t). \quad (3)$$

Since Equation 2 is non-linear an extended Kalman filter (EKF) is needed. The EKF equations presented here are based on those of Welch and Bishop [12].

In order to apply the EKF we first need to estimate the uncertainty in the model Equations 1 and 3. There are three main causes of uncertainty in this system:

1. The assumption that the motion parameters u, v and ω are constant may not be true. This introduces error terms $\varepsilon_u, \varepsilon_v$, and ε_ω with variances σ_u^2, σ_v^2 and σ_ω^2. These errors affect all lines equally.

2. Each line may have an independent deviation from the motion model. This allows for deformations in the appearance of the object not accounted for by rigid motion within the image plane. For the ith line the errors in ρ and θ are denoted ε_{ρ_i} and ε_{θ_i}, and have variances σ_{ρ_i} and σ_{θ_i}.

3. The measurements of ρ and θ in Equation 3 are uncertain due to the quantisation of the Hough accumulator. For the ith line these errors are e_{ρ_i} and e_{θ_i} with variances s_{ρ_i} and s_{θ_i}. If each accumulator cell is $\Delta_\rho \times \Delta_\theta$ then $s_{\rho_i} = \Delta_\rho^2/12$ and $s_{\theta_i} = \Delta_\theta^2/12$.

This gives two error vectors associated with the EKF:

$$w_t = [\varepsilon_u, \varepsilon_v, \varepsilon_\omega, \varepsilon_{\rho_1}, \varepsilon_{\theta_1}, \dots, \varepsilon_{\rho_n}, \varepsilon_{\theta_n}]^T, \text{and} \quad (4)$$

$$v_t = [e_{\rho_1}, e_{\theta_1}, \dots, e_{\rho_n}, e_{\theta_n}]^T. \quad (5)$$

All of these error terms are assumed to be independent, zero-mean, and Gaussian. ε_u, ε_v, ε_ω, ε_{ρ_i}, and ε_{θ_i} add to all occurrences of u, v, ω, ρ_i, and θ_i in Equations 1 and 2. The terms e_{ρ_i} and e_{θ_i} add to ρ_i and θ_i in Equation 3.

From these model equations and error terms we form the EKF as

$$\tilde{s}_t^- = f(\tilde{s}_{t-1}) \quad (6)$$

$$P_t^- = A_{t-1} P_{t-1} A_{t-1}^T + W_{t-1} Q_{t-1} W_{t-1}^T \quad (7)$$

$$K_t = P_t^- H_t^T (H_t P_t^- H_t^T + V_t R_t V_t^T)^{-1} \quad (8)$$

$$\tilde{s}_t = \tilde{s}_t^- + K(m_t - h(\tilde{s}_t^-)) \quad (9)$$

$$P_t = (I - K_t H_t) P_t^-, \quad (10)$$

where $\tilde{s}_t$ is the estimate of the state at time t and P_t is the covariance of this estimate. $\tilde{s}_t^-$ and P_t^- are the initial estimates made from the model at time $t-1$, and K_t is the Kalman gain matrix used to combine this initial estimate with the measurement. The remaining terms, A_t, W_t, H_t and V_t are Jacobian matrices defined as

$$A_{[i,j]} = \frac{\partial f_{[i]}}{\partial s_{[j]}}(\tilde{s}_t), \qquad W_{[i,j]} = \frac{\partial f_{[i]}}{\partial w_{[j]}}(\tilde{s}_t), \tag{11}$$

$$H_{[i,j]} = \frac{\partial h_{[i]}}{\partial s_{[j]}}(\tilde{s}_t^-), \qquad V_{[i,j]} = \frac{\partial h_{[i]}}{\partial v_{[j]}}(\tilde{s}_t^-). \tag{12}$$

There is one final complication to consider in tracking lines. The θ values for a line are restricted to the range $[0, \pi]$. A moving line can cross the ends of this range, and its measurement will be returned mod π. Cycling around the range of θ introduces a sign change in ρ. Both of these effects must be carefully accounted for in the Kalman filter implementation and line finding routines.

4 A Combined Kalman Filter and Hough Transform

The Kalman filter provides an estimate of a line's location in each frame. This can be used to restrict the search for lines in the Hough space, and so reduce the computational load. As well as an estimate, $(\tilde{\rho}, \tilde{\theta})$, of a line's location, the Kalman filter provides the variances, σ_ρ^2 and σ_θ^2, in these estimates. Using this information we restrict the search for a line in the Hough space to values of ρ and θ in the ranges $(\tilde{\rho} - k\sigma_\rho, \tilde{\rho} + k\sigma_\rho)$ and $(\tilde{\theta} - k\sigma_\theta, \tilde{\theta} + k\sigma_\theta)$, where k is some constant. In the experiments presented here we have used $k = 2$, which corresponds to a 95% confidence interval.

A naive implementation to compute an $n \times n$ Hough accumulator from an $n \times n$ image is $O(n^3)$. If the values of ρ and θ are reduced to a range of $m \ll n$ values, then this drops to $O(nm^2)$, as we can iterate over one of the image coordinate axes and then over the m ρ and θ values, solving for the other image coordinate each time. Once the Kalman filter has converged, m is small (in our experiments, less than 10) and can be treated as a constant. This means that in practice, tracked lines can be found in $\sim O(n)$ time.

In order to evaluate the efficiency of the proposed techniques a series of experiments have been conducted on synthetic images. These evaluate the method's robustness to image noise, occlusion, and violations of the constant motion assumption. A simple image sequence was constructed, consisting of a square moving against a uniform background. The true parameters of the four edges of the square are compared with the prediction provided by the Kalman filter to measure its effectiveness.

In order to determine the effects of image noise, Gaussian noise with a range of variances was added to the input images. The range of intensity values was from 0 to 255, and noise with a standard deviation of up to 50 intensity levels was found to have only a small effect on the results, as shown in Figure 1. This is due to the fact that the Hough transform finds features in a global rather than local fashion, so the effects of image noise are suppressed in the Hough space.

The second factor that we consider is occlusion. As with image noise, the use of a Hough transform means that the method is quite robust to partial occlusion. The input images for this experiment had circular blobs placed over the centre of each side of the square. The radius of these occluding regions was varied to give a different levels of occlusion. As Figure 2 shows, up to around 70% occlusion is handled by the system with only small increases in the errors of the technique.

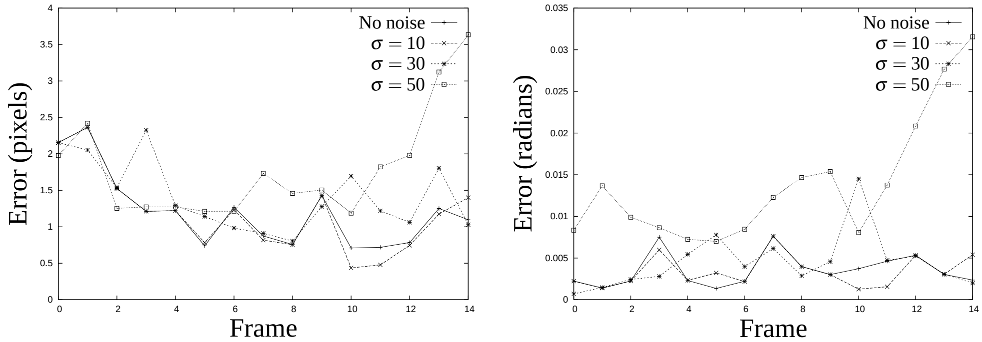

Figure 1: Errors in the EKF predictions of ρ (left) and θ (right) under varying levels of image noise. The image noise is zero mean Gaussian noise with variance σ^2.

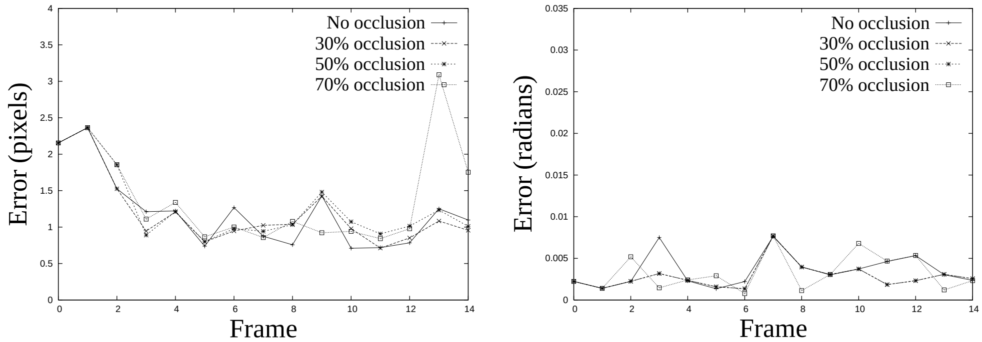

Figure 2: Errors in the EKF predictions of ρ (left) and θ (right) under varying levels of occlusion.

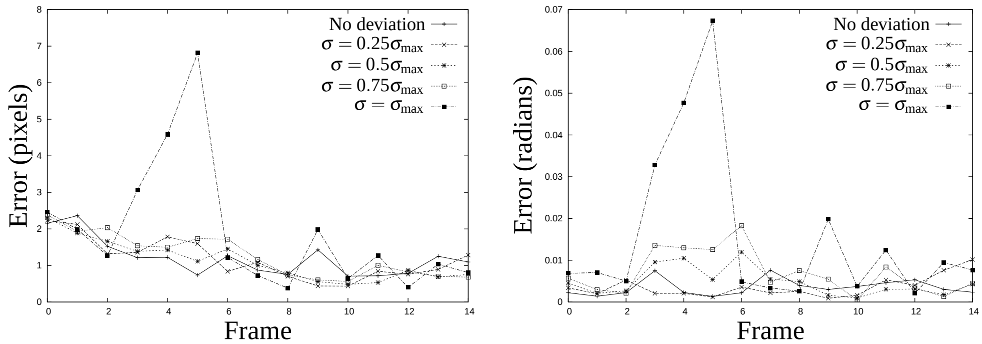

Figure 3: Errors in the EKF predictions of ρ (left) and θ (right) under varying levels of deviation from the constant motion model. Deviation from the model is expressed as a multiple of the variance of the expected deviation, $\sigma_{\max}^2$.

Next we consider deviations from the constant motion model. As discussed in Section 3 estimates of the level of deviation from the model are provided to the EKF. In order to evaluate the effectiveness of this error model noise was added to the motion parameters of the square. This noise was Gaussian, and different variances were used. These variances were expressed as multiples of the estimated variances of $\varepsilon_u, \varepsilon_v$, and ε_ω. Figure 3 shows the results of this experiment. When the error in the motion has a variance less than that expected by the model there is little change in the accuracy of the results.

The time taken to compute the Hough transform using the integrated approach was evaluated on real rather than synthetic images. This is because real images contain more edge points, and ground truth motion is not required to time the algorithm. Figure 4 shows the time taken to compute the Hough transforms needed for a real image sequence (the 'square' sequence to be presented in Section 5). The integration of the Kalman filter and Hough transform greatly reduces the computation needed to detect lines. The filter rapidly converges, and requires only about 0.5 seconds to detect lines in each frame. The implementation of the Hough transform has no other optimisations applied to it, and requires over a minute for to compute a full accumulator array at the same resolution.

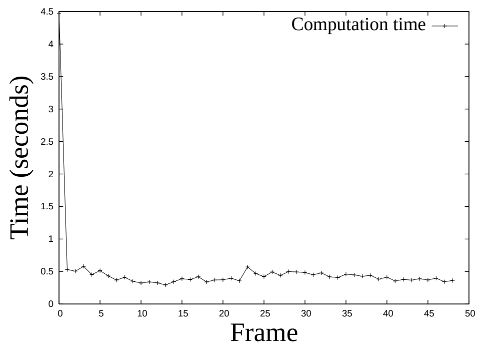

Figure 4: Computation time for the Hough transform in the 'square' sequence

5 Results

Figure 5 shows results of the Hough-Kalman tracking algorithm on several real sequences. In each case several strong lines are identified manually in the first frame, then tracking proceeds automatically. Each of these sequences has particular problems, which illustrate the strengths and weaknesses of our technique.

In the first sequence a square object is manually rotated. In this sequence the motion is somewhat irregular, with pauses in the motion as the grip on the object is changed, and there are perspective distortions as the square is rotated in depth. Although neither of these problems is explicitly accounted for in the Kalman filter model equations, using wider distributions for the error terms accounts for their effects. This does, however, mean that the predicted and measured location of the square do not always correspond, as in the second frame shown. The Kalman filter does not, however, aim to model the exact location of the object in each frame. Rather it builds an overall model of the motion which best fits the observed sequence as a whole.

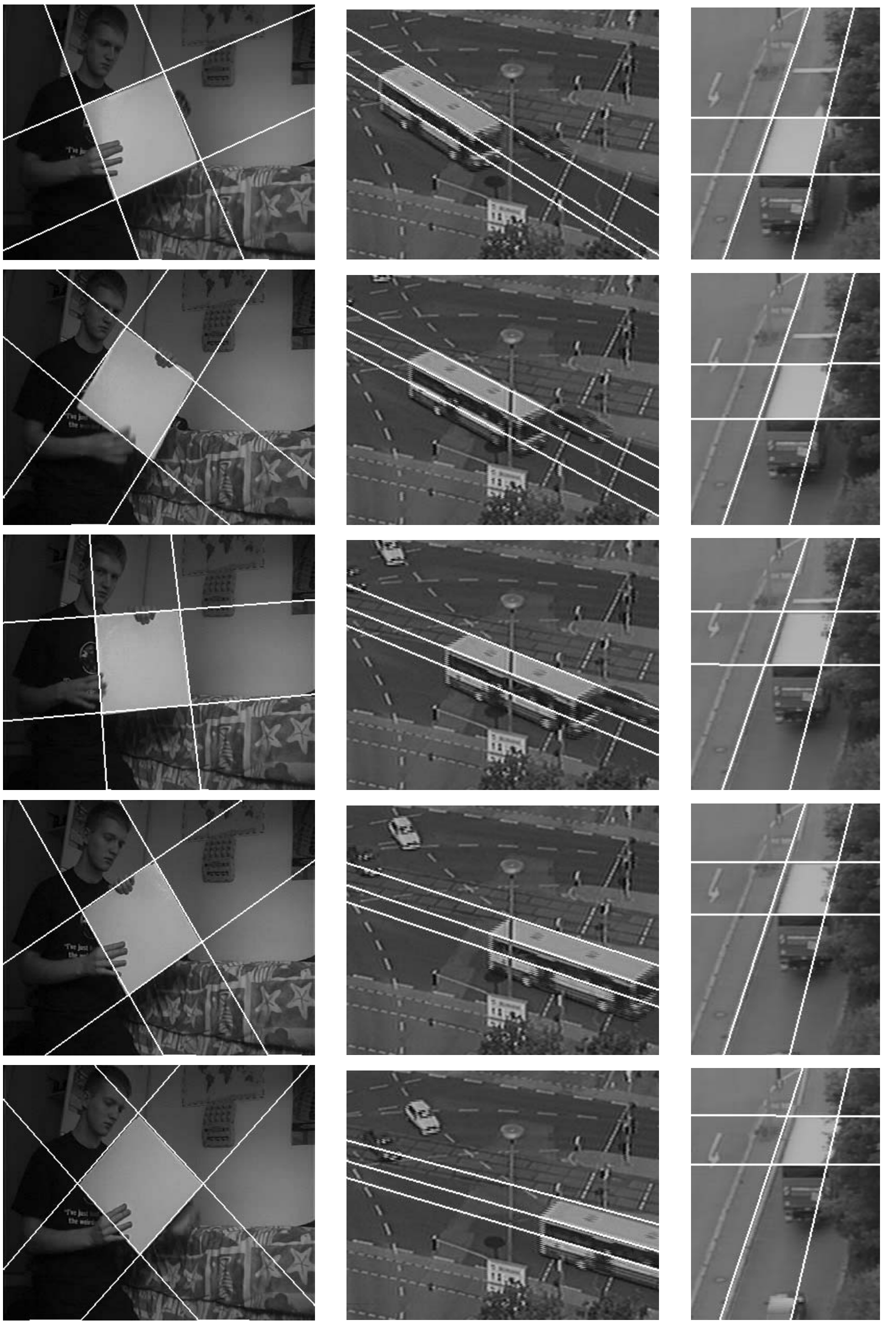

Figure 5: Tracking in a real video sequence. The left column shows every 10th frame tracking a square object, the centre column every 10th frame tracking a bus, and the right every 15th frame tracking a truck.

The second sequence shows the tracking of a bus as it turns a corner. In this case there are several strong lines running along the bus. Due to the shape of the bus and interlacing effects, however, there are no strong vertical lines on the object. This introduces an ambiguity in the motion – essentially there is an aperture problem, where motion along the direction of these lines is difficult to determine. Despite these problems, the lines are successfully tracked over 50 frames.

The final sequence shows the tracking of a truck as it moves along the road. There are two main problems here, associated with distraction and occlusion. The first issue is that the centre road markings form a strong line near to the left hand side of the truck. This line is chosen by preference, as it has a stronger peak in the Hough transform. Since the motion of the truck is parallel to this line, it does not adversely affect the tracker. Lines which are not close to those on the object in the Hough space are ignored. Those which lie at a significant angle to the motion, such as the road marking in front of the truck in the third frame shown, are either not chosen or do not persist as they conflict with the Kalman filter's motion model.

The second problem in this sequence is occlusion caused by the trees on the right of the frame. These obscure the right hand side of the truck. Initially the line is tracked successfully, as in the fourth frame shown, using the estimates from the Kalman filter based on the other three lines' motion. During the occlusion, however, a distracting line is introduced as a second vehicle enters the frame. This introduces a line with similar orientation and location to the occluded one. Since the motion of this line is coincidentally consistent with the Kalman filter it is tracked as if it were part of the original object. Two possible extensions to our method could reduce this problem.

The first is to try and evaluate the tracking results in order to determine if a measurement is trustworthy or not. Measurements that are deemed untrustworthy can be disregarded, and the prediction from the Kalman filter used instead. This decision could be made by comparing the motion of each line to that of other lines on the object. Lines whose motion does not fit with the object as a whole can be disregarded. Occluded lines are particularly sensitive to strong distractors, and can be identified because the response of a line in the Hough accumulator weakens as it becomes occluded.

A second approach that could reduce the problem of distracting lines is the use of local processing. Since the Hough transform is a global approach it gathers evidence for lines from the whole image. A moving object, however, typically only occupies a small region of the image. If this region can be identified then points outside that region can be excluded from consideration. This would reduce the effects of distractions caused by distant objects, such as the car that enters the 'truck' sequence in Figure 5. It would also further reduce the computational load of the Hough transform, as the image size would effectively be reduced to the size of the object being tracked. The Kalman filter's motion estimates could be used to continually update the image region considered.

6 Conclusion

In this paper we have presented a method for tracking that combines the Hough transform and Kalman filter. Experiments on synthetic images show that this method is robust to image noise, partial occlusions, and deviations from the constant motion model. Strong distracting lines can cause problems, particularly when combined with occlusion. The

Kalman filter predictions are used to reduce the computational load of the Hough transform for lines from $O(n^3)$ to $\sim O(n)$, resulting in dramatically reduced computation times.

The method presented here is based on tracking lines with a constant motion model. Other features and motions could be easily accounted for within the same framework. The Hough transform has successfully been applied to a range of image features, and the Kalman filter can be modified to allow for other motion models.

Acknowledgements

The traffic image sequences used in Figure 5 are from the KOGS/IAKS image sequence server at Universität Karlsruhe [7].

References

[1] T. Behrens, K. Rohr, and H. S. Stiehl. Using an extended Hough transform combined with a Kalman filter to segment tubular structures in 3D medical images. In *Vision, Modelling, and Visualisation*, 2001.

[2] R. O. Duda and P. E. Hart. Use of the Hough transform to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15, 1972.

[3] G. L. Foresti. A line segment based approach for 3D motion estimation and tracking of multiple objects. *International Journal of Pattern Recognition and Artificial Intelligence*, 12(6):881–900, 1998.

[4] M. Hills, T. Pridmore, and S. Mills. Object tracking through a Hough space. In *Proceedings of Visual Information Engineering*, pages 53–56, 2003.

[5] B. K. P. Horn and B. G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[6] P. V. C. Hough. Method and means for recognising complex patterns. U.S. Patent 3,069,654, 1960.

[7] Institut für algorithmen und kognitive systeme. [online], available from `http://i21www.ira.uka.de/image_sequences/index.html`. Last accessed 7 April 2003.

[8] R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME – Journal of Basic Engineering*, 82(Series D):35–35, 1960.

[9] J. M. Nash, J. N. Carter, and M. S. Nixon. Dynamic feature extraction via the velocity Hough transform. *Pattern Recognition Letters*, 18(10):1035–1047, 1997.

[10] P. Nesi, A. Del Bimbo, and D. Ben-Tzvi. A robust algorithm for optical flow estimation. *Computer Vision and Image Understanding*, 62(1):59–68, 1995.

[11] R. S. Stephens. A probabalistic approach to the Hough transform. In *Proceedings of the British Machine Vision Conference*, pages 55–60, 1990.

[12] G. Welch and G. Bishop. An introduction to the Kalman filter. Technical Report TR95-041, Department of Computer Science, University of North Carolina, 1995.