Detection and Tracking of Very Small Low Contrast Objects

D. Davies†, P. Palmer†, and M. Mirmehdi‡

†School of Electrical Engineering, Information Theory & Mathematics,
University of Surrey, Guildford GU2 5XH, England

‡Computer Vision Group, Dept. of Computer Science,
University of Bristol, Bristol BS8 1UB, England

Abstract

We present a Kalman tracking algorithm that can track a number of very small, low contrast objects through an image sequence taken from a static camera. The issues that we have addressed to achieve this are twofold. Firstly, the detection of small objects comprising a few pixels only, moving slowly in the image, and secondly, tracking of multiple small targets even though they may be lost either through occlusion or in noisy signal. The approach uses a combination of wavelet filtering for detection with an interest operator for testing multiple target hypotheses based within the framework of a Kalman tracker. We demonstrate the robustness of the approach to occlusion and for multiple targets.

1 Introduction

One of the most difficult goals of Automatic Target Recognition (ATR) is to spot incoming targets at long range where the motion is small and signal to noise is poor, and to be able to track such targets long enough to identify whether the target is approaching in an aggressive manner. In [9], some of the current authors determined the best magnification to zoom in onto a target object while it is still at a long range. However, the important first-step issue for the detection of the very small target itself was not addressed. We are interested in locating long-range moving objects in FLIR images where the object may only be a few pixels in size and has low contrast with its background. Simple methods of image differencing [11] are therefore not appropriate. We need a more robust algorithm that is fairly simple to compute, has the potential to run in real-time, and is relatively immune to noise. We emphasise that template matching methods are inappropriate to this problem as our targets are very small.

This problem has been addressed by Blostein and Huang [1] who used multistage hypothesis testing on a large number of hypothesised trajectories. Their approach looks for changes in mean pixel value through the sequence, but therefore requires trajectory testing on a per pixel basis. The authors have argued that there is a need for an effective decision theoretic approach to the detection of small low-contrast objects. In this paper

BMVC 1998 doi:10.5244/C.12.60

we expand upon the ideas of [1], exploiting useful properties of wavelet filters to provide detection of the motion of these small, low-contrast objects.

An alternative approach has been to try to improve the signal to noise ratio by removing clutter from the image. This approach has been described by Ffrench et al. [6] using a 2-D adaptive lattice algorithm for the removal of correlated clutter. Comparisons with mainly simulated images showed that while better results are obtained, the increased cost in computational time and algorithm complexity is considerable.

In this paper we exploit useful properties of wavelet filters [5, 3, 12] to provide a robust method for object detection even when the object size is as small as 6 pixels. We anticipate that the object size and 2D motion will change during the sequence, as the target approaches. We have developed a high order prediction mechanism for use with the Kalman filter [2] to determine regions of interest which reduces the search space for generating hypothesised tracks. The method generates a hypothesis tree for the motion of target objects across the image sequence which is pruned using a *measure of interest* operator based upon the Kalman filter error estimates. The interest operator is built upon the principles described in [4].

The novelty of our approach is that we demonstrate detection of moving objects even before they become apparent to the naked eye in the image sequence; we successfully track these objects long enough to estimate their trajectories. We have a multiple target tracker which enables us to track many targets simultaneously, which enables us to pick up targets after passing through occlusion and to continue to track them. The processing time for this algorithm is typically below 1 second per frame without any emphasis on optimisation of computational time, and the method is highly parallelisable making it amenable to real time implementation.

In Section 2 we describe the wavelet filtering that detects small low contrast targets. In Section 3 we describe the Kalman tracker methodology used for multiple target tracking with automatic addition and removal of targets. In Section 4 we discuss the hypothesis tree and the interest operator. Finally we present our results in Section 5 and summarise and conclude in Section 6.

2 Detecting Small Targets using Wavelets

Our emphasis is on the detection of small targets only a few pixels across. For such objects a wavelet description is most suitable as wavelets encapsulate information on discontinuities in a very succinct way. We could describe the object in terms of a wavelet decomposition and then apply a matched filter to it. However, the objects we seek have very little structure associated with them. The approach we have adopted uses a *temporal* filtering of the image sequence using a wavelet filter rather than a spatial filtering across the image [13].

The classic way to detect moving objects in an image sequence taken with a fixed camera is by differencing neighbouring images. The problem with this approach is that many noise variations are also detected, and if our object is small it can prove hard to separate out the moving target from the noise. We may consider the differencing as applying a Haar wavelet filter temporally at a fixed pixel location. Therefore image differencing is a temporal wavelet filter using the smallest known filter size. Use of larger wavelet filters then provides a way of combining multiple images [11].

Wavelet filters have been proven to be superior than Fourier methods for detecting localised high frequency behaviour in an otherwise predominantly low frequency signal [5, 13, 7, 8]. This property stems from the fact that they are defined as scaled and dilated copies of a single function $\psi(x)$ having the property:

$$\int_{-\infty}^{\infty} \psi(x) dx = 0 \tag{1}$$

The wavelet functions are then defined by:

$$\psi_{mn}(x) = 2^{-m}\psi(2^{-m}x - n) \tag{2}$$

where m is the dilation factor and n represents the shift. Wavelets are associated with a *scaling* function $\phi(x)$ via the *two-scale equation*:

$$\phi(x) = \sum_n h_n \phi(2x - n) \tag{3}$$
$$\psi(x) = \sum_n g_n \phi(2x - n), \tag{4}$$

The finite set of coefficients h_n and g_n constitute low pass and high pass filters respectively [5]. For our purposes we have used the Daubechies wavelet filters g_n of lengths 4, 6 and 8. These filters then provide the generalisation of image differencing to combining 4, 6 and 8 images in a sequence together. The high pass nature of the filter then detects the passage of a greylevel boundary across each pixel and so highlights the outline of the object. This is illustrated in Figure 1.

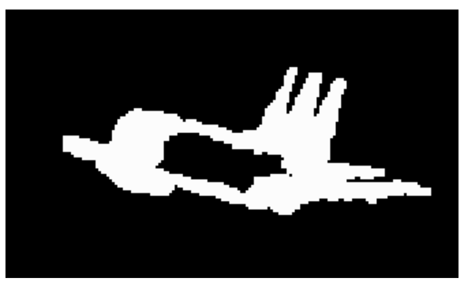


Figure 1: Output of the temporal wavelet filter using 4 frames. This shows the motion of an airplane and how the temporal filter captures boundary motion. The body of the airplane is not found at all.

The advantage of using these wavelet filters is that they are robust to noise. By combining evidence over several frames coherent motion stands out above the noise, providing a much higher signal to noise ratio.

For each filtered image we compute a channel energy in local neighbourhoods by averaging the absolute values of the wavelet coefficients. This is then thresholded and passed through a morphological opening process (erosion followed by dilation) which removes isolated noise regions leaving just the targets and one or two larger noise regions. A connected component analysis is then performed to produce a set of connected regions, which is then passed to the Kalman filter.

3 Multiple Target Kalman Tracker

Having detected small objects moving slowly across the image, we track the targets using a Kalman tracker. There will still persist a few random noise targets appearing among detected objects, so an important part of the Kalman tracking process is to accurately predict the location of each target in the next frame in order to minimise the number of false matches and reduce the number of hypothesised trajectories. Since the targets are small we only track their centres of mass.

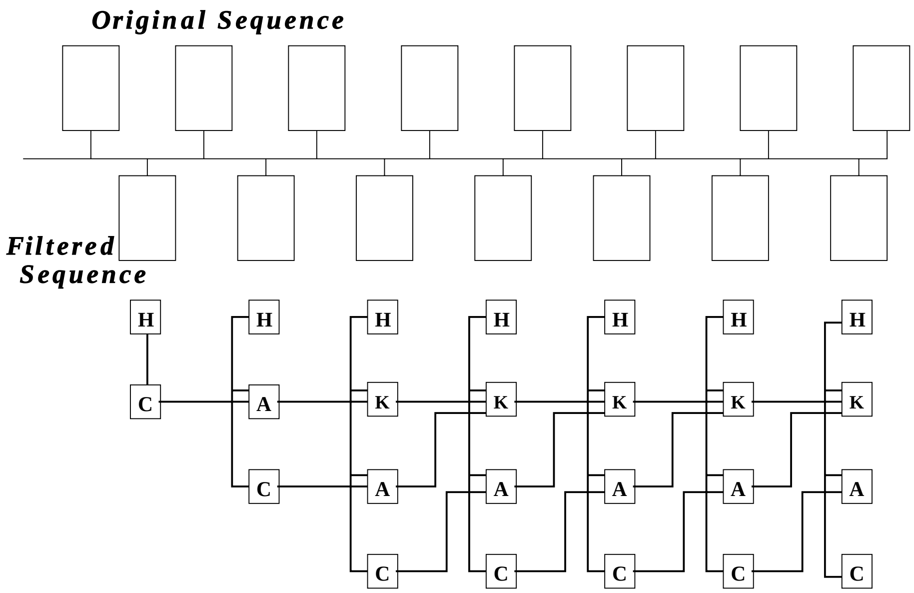


Figure 2: Schematic representation of the Kalman tracker system. For each filtered image we obtain a list of 'hopefuls' (H) from the wavelet filter. We then try to match these objects with the list of Kalman tracked objects (K). The remaining hopefuls are then matched with associations (A) in the previous two images, and the rest are left in a list of candidates (C).

The Kalman state variable in our method comprises of the centres of mass over three successive frames. We employ a constant acceleration model with state transition matrix:

$$\Phi = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 3 \end{pmatrix} \tag{5}$$

This predicts the location of the target in the next frame to fourth order in the frame interval. In the ideal case, combining the observations over four frames with the weightings given by the above matrix should give zero. We therefore combine the observations over four frames and use the deviation from zero as the measure of system noise:

$$\omega_{n+2} = z_{n+2} - 3z_{n+1} + 3z_n - z_{n-1} \tag{6}$$

where z_n is the observed position in frame n.

We employ an interest operator to produce the matches between targets in successive frames based around the Kalman prediction. Some comment is necessary to describe the initialisation of this procedure. We start by detecting candidate objects (C) in the first frame, using the wavelet filtered images. We then seek targets in the second frame in the vicinity of the previous location. Once such an association has been made we identify this candidate as an *association* (A). In the third frame we use a constant velocity (linear) prediction for each association. The interest operator is then used to identify whether any candidates in this frame match the prediction. We describe the interest operator below. If we can follow the target across three successive frames we then initialise a new Kalman tracker for this target (K). The system is described by Figure 2.

At present once an object is lost, either disappearing into the noise or through occlusion that Kalman tracker terminates and the list of objects is reduced by one. It would be a straight-forward extension to allow the prediction to persist over more frames. An alternative strategy that we may consider is to combine tracked segments together in a higher level algorithm searching for potentially aggressive motion. This would prove a more robust algorithm as it combines evidence over a larger time interval when identifying target trajectories.

4 Hypothesis Testing - the Interest Operator

The use of the wavelet transform combined with the Kalman filter provides a very robust method for tracking small targets. Even so, each triplet of targets in successive frames will generate a new initialisation of a Kalman tracker. It is therefore essential that the number of hypothesised trajectories is maintained at a comparable level to the number of Kalman tracked objects.

To achieve this we have used an interest operator using the same design principles as that in [4] to prune the search tree. This operator is based upon the error covariance generated by the Kalman filter, to predict anticipated errors in the measurements. If the matrix H maps the state space into measurement space, $\hat{x}^-$ is the apriori state estimate, and P^- is the normal error covariance in the state prediction, then the expectation of $\mathcal{E}(\mathbf{z}-H\hat{\mathbf{x}}^-)(\mathbf{z}-H\hat{\mathbf{x}}^-)^T$ is HP^-H^T. This matrix is updated during the standard Kalman update cycle, and thus provides estimates of the region around each prediction where we may expect to find the measurement (σ_x, σ_y). Since the motion of these targets is small, the predictions are fairly accurate and so this effectively reduces the search space. The interest operator is defined as:

$$I(\Delta_x, \Delta_y) = M\mathrm{e}^{-K} \tag{7}$$

where

$$K = -\alpha\sqrt{\left(\frac{\Delta_x^2}{2\sigma_x^2} + \frac{\Delta_y^2}{2\sigma_y^2}\right)}$$

and M is the number of pixels associated with the candidate from the connected component analysis. The terms Δ_x, Δ_y represent the differences in the x and y coordinates between the target's centre of mass as predicted by the Kalman tracker and the observed centre of mass of a candidate object. The constant parameter α is a measure of how far away from the prediction a particular measurement may be to still hold some interest.

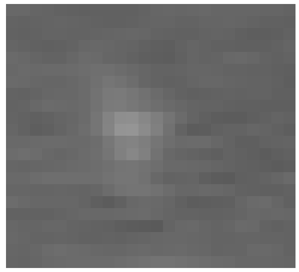


Figure 3: Profile of aircraft in frame 32 of the sequence. It is only 6-7 pixels across throughout the early part of the sequence.

This has been fixed at 2 standard deviations. The interest is biased towards larger candidates that fit the prediction through the measure M. This also takes account of the fact that many targets can be expected to become larger as they approach.

The interest operator guarantees that virtually all hypotheses apart from the correct one have a very low interest level and are quickly rejected. The longer the object is tracked well the more robust the interest operator becomes in identifying the true object in the presence of other close and/or high contrast decoys, because the error behaviour itself is modelled more consistently.

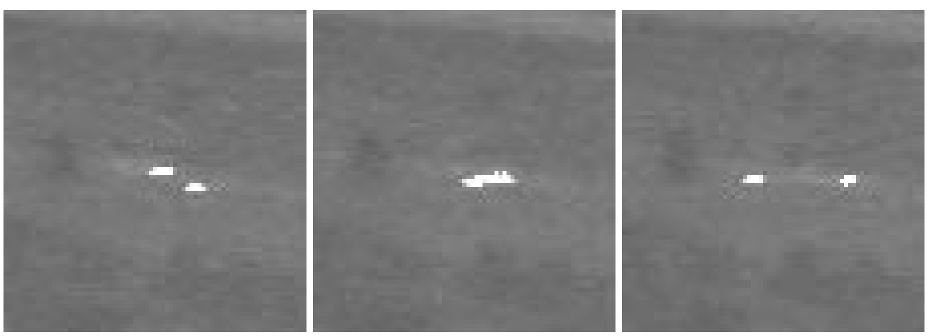
5 Experimental Results

We have used our algorithm on a sequence of FLIR images taken by a fixed sensor looking out over a plain. In the sequence an aircraft is approaching and there is a road in the middle distance with cars travelling along the road. Our aim is to try and detect all these moving objects and to track them at least long enough to determine whether their trajectories are hostile. All the objects are low contrast and are hard to pick out in the image sequence. In Figure 3, we show an enhancement of the aircraft from frame 32 of the sequence. We show results from our tracker algorithm in Figure 4 which shows the scene from the camera and gives an impression of the size of the targets being tracked. In this image the airplane appears near the top left of the image, and we are tracking 3 cars along the road.

The tracking of the cars along the road is more intermittent than the aircraft due to occlusion by bushes and hedges along the road. Near the start of the sequence, however, we manage to track two cars which approach each other and pass by. We successfully track these two cars through approach (in Figure 5(a)), on passing (5(b)), and on separation (5(c)).

Towards the end of the sequence the aircraft passes overhead and is quite large. The tracker follows the centre of mass, but by using the connected regions we have then computed a correlation with the greylevels themselves to produce the displayed images. The connected regions, however, produce an accurate description of the actual shape of the target. Although this cannot be seen in the early part of the sequence, in Figure 6 we show the connected region superimposed on frame 65 of the aircraft, and alongside the original greylevel image. This shows how well the connected region defines the shape and

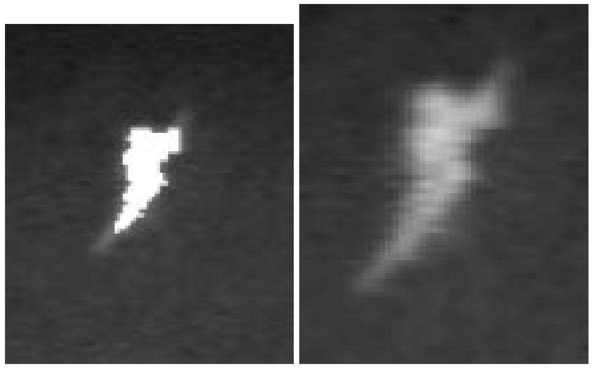
Figure 4: Multiple objects: 1 airborne craft and 3 vehicles



(a) (b) (c)

Figure 5: Two approaching cars approaching,passing and separating

hence orientation of the target as it approaches. This may also provide useful information in determining target pose - an important ingredient in determining whether it is hostile.



(a) Connected Region (b) Original Image (Enhanced)

Figure 6: Intensity profile of target at close range (b) and the connected region obtained from the wavelet transform following morphology (a).

To get an idea of how well the targets are tracked across the whole sequence we have produced an MPEG movie, downloadable from the WWW site:

http://www.ee.surrey.ac.uk/EE/VSSP/demos/smallobj/index.html

Only by this means can a full appreciation of the tracker be made. In order to give some idea of performance we have combined the targets throughout the whole sequence and superimposed them onto the last image. This is shown in Figure 7 which gives a clear idea of how the aircraft and cars moved through the sequence and how well we managed to track them. This illustrates the speed of the aircraft relative to the cars as the aircraft trajectory marked is disjoint.

6 Conclusions

Our algorithm combines the discriminating power of the wavelet transform with a Kalman filter to *detect* and *track* the motion of small, low contrast objects in image sequences. We have demonstrated the algorithm using a sequence of FLIR images which incorporate slow and fast motion of an approaching aircraft as well as the semi-hidden motion of cars along a country road. We have shown that we can track a substantial part of this motion even though the targets have very poor signal to noise and may become occluded.

The method we have described is highly parallelisable since we filter each pixel in the image sequence separately, and the multiple Kalman targets are all independent. Although this has not been parallelised in any way, the computational time for our sequence of 75 images was around 3.5 minutes on a Ultra SPARC 450 machine. This corresponds to 2.5 seconds per frame.

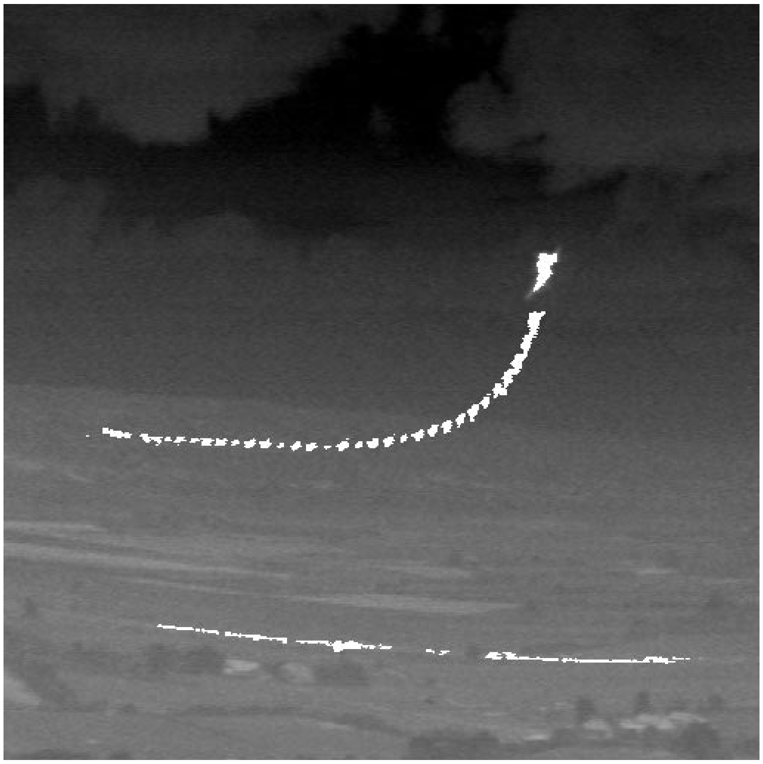


Figure 7: The whole sequence of tracked objects superimposed upon the last image of the sequence to give an idea of the motion and shift in target sizes.

The main drawback of the method at present is that for longer wavelet filters the targets become smeared over a larger region of the images and so locality of the targets is lost. We have overcome this deficiency by performing a correlation analysis with the image greylevels when generating the results. This does not, however, always do justice to the tracker algorithm. In the cases presented, our targets are always reasonably small and so the problem has not been too important.

We believe that the method is robust in that the target is always tracked even when decoys are employed. In order not to be overcome by too many trackers active at the same time we have introduced an interest operator that prunes the tree of hypothesised trajectories.

The method also provides a 'shape' of a region enclosing the target object, which provides information about the manoeuvring of the aircraft and its change in pose. Shape characterisation could be achieved by utilising higher moment information about the target's centre of mass. Shape information could conceivably be incorporated with the interest operator function, or retained for further analysis by a subsequent pattern recognition device.

The method can be combined with noise reduction algorithms and/or in conjunction with the control mechanism of the tracking IR camera. For example, it could be used to zoom onto the target [9] to obtain more detailed information to complete the object recognition process. It would therefore also be part of the overall control of the vision process [10].

References

[1] S. Blostein and T. Huang. Detecting small, moving objects in image sequences using sequential hypothesis testing. *IEEE Transactions on Signal Processing*, 39(7):1611–1629, 1991.

[2] R.G. Brown and P.Y.C Hwang. *Introduction to Random Signals and Applied Kalman Filtering*. John Wiley & Sons, 1985.

[3] C.K. Chui. *An Introduction to Wavelets*. Academic Press Inc, 1992.

[4] H. Dabis, P. Palmer, and J. Kittler. An interest operator based on perceptual grouping. In *Proceedings of the Scandinavian Conference in Image Analysis*, pages 315–322, 1994.

[5] Ingrid Daubechies. *Ten Lectures on Wavelets*. SIAM, 1992.

[6] P.A. Ffrench, J. R. Zeidler, and W. H. Ku. Enhanced detectability of small objects in correlated clutter using an improved 2-d adaptive lattice algorithm. *IEEE Trans. on IP*, 6(3):383–397, 1997.

[7] I.Daubechies. Orthonormal bases of compactly supported wavelets. *Communications on Pure Applied Mathematics*, 41:909–996, 1988.

[8] S. Mallat and W.L. Hwang. Singularity detection and processing with wavelets. *IEEE Transactions on Information Theory*, 38(2):617–641, March 1992.

[9] M. Mirmehdi, P.L. Palmer, and J. Kittler. Towards optimal zoom for automatic target recognition. In *Proceedings of the Tenth Scandinavian Conference in Image Analysis, Volume I*, pages 447–453, 1997.

[10] M. Mirmehdi, P.L. Palmer, J. Kittler, and H. Dabis. Complex feedback strategies for hypothesis generation and verification. In *Proceedings of the 7th British Machine Vision Conference*, pages 123–132. BMVA Press, 1996.

[11] P.L Rosin and T. Ellis. Image difference threshold strategies and shadow detection. In *Proceedings of the 6th British Machine Vision Conference*, pages 347–356. BMVA Press, 1995.

[12] G. Strang and T. Nguyen. *Wavelets and Filter Banks*. Wellesley Cambridge, 1996.

[13] R.N. Strickland and H.I. Hahn. Wavelet transform methods for object detection and recovery. *IEEE Transactions on Image Processing*, 6(5):724–735, May 1997.