Robust Tracker of Small, Fast-moving Low-contrast Targets

D. Davies[1], P. Palmer[1], & M. Mirmehdi[2]
[1]School of Electrical Engineering, Information Theory & Mathematics,
University of Surrey, Guildford GU2 5XH, England

[2]Computer Vision Group, Dept. of Computer Science,
University of Bristol, Bristol BS8 1UB, England

ABSTRACT

We present a multiresolution adaptive wavelet transform to locate small low-contrast targets. Our approach expands upon methods which use adaptive filters to remove noise to produce a near real-time robust tracker using a specially adapted Kalman filter. This generates a small set of hypotheses to test. Incorrect hypotheses are removed using an interest operator founded on the error covariance generated by the Kalman filter. We demonstrate the technique using some experimental results.

1 INTRODUCTION

We are interested in detecting 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 are therefore not appropriate. We need a more robust algorithm to correctly detect the object while it is still at long range, and yet be simple to compute and be relatively immune to noise.

Blostein and Huang [2] used sequential hypothesis testing to examine candidate object trajectories and tested sequentially for a shift in mean intensity in image sequences. However, the technique requires hypothesis testing and the examination of every pixel in every frame of a sequence. Ffrench et al. [6] have developed an improved 2-D adaptive lattice algorithm geared towards ameliorating the image through the removal of correlated clutter to enhance the detectability of small objects. The 2-D adaptive lattice algorithm generally removes the clutter such that the hypothesised object of interest is left (depicting a stronger signal). They compare their filters against 2-D LMS filters for clutter removal in mainly simulated images, and they conclude that while they obtain better results, the increased cost in computational and algorithm complexity is considerable.

The authors of [2] have argued that there is a need for an effective decision theoretic approach to the detection of small low-contrast objects. This paper attempts to address this issue. We combine the use of the fast wavelet transform for adaptive filtering of images which provides a multiscale representation for object detection. We detect objects of sizes in the order of 10 pixels, although the object size may, and usually would be expected to change during the sequence. A Kalman filter is used to determine regions of interest which further reduces the computational time for searching. The method generates a hypothesis tree for the motion of the target object across the image sequence which is pruned using an interest operator based upon the Kalman filter error estimates. The whole process could act as a front-end to a comprehensive target detection and recognition system.

We demonstrate in this paper the successful tracking of small objects, even with decoys in close proximity.

2 THE ALGORITHM

Wavelets are sets of orthogonal functions which have limited spatial and frequency content and provide a multiresolution decomposition of a signal. Consequently they are ideally designed to locate target objects of small size in images.

The wavelet transformation of an image produces four lower resolution images comprising of a coarser scale space representation of the image and three difference signals (see [10] for details). This decomposition has proved useful for a number of applications of image processing ([7]). For small objects we are particularly interested in high pass filters, but the most interesting channel for our application is a high pass filter in the vertical direction and a low pass in the horizontal. This allows us to reduce the effects of noise while still seeking small scale rapid variation in greylevel in the direction of principle change. We note that the same wavelet space was used in [1] for detecting tanks in a rough terrain scene.

From the wavelet transformation (WT) of the image we compute energies in each channel and threshold to identify possible targets. These energies are defined in windows which are specified by the scale of the target object we are seeking. We assume this is known a priori. The wavelet transformation is therefore carried out in overlapping windows and windows of high transform energy are merged into a region which should cover the target object.

Once targets have been found, using the energy method described above, an estimate of the motion is made based upon nearest neighbours in the subsequent frames. Any hypothesised target that is insufficiently significant is rejected. Once potential targets have been tracked across three successive frames a Kalman filter (see [3]) is initialised to predict the subsequent motion. The Kalman model has been especially adapted to produce reliable estimates of the motion, so that false alarms can be excluded at a very early stage. The use of the wavelet transform combined with the Kalman filter provides a very robust method for removing inaccurate models. Even so, an interest operator is introduced using the same design principles as that in [4] to prune the search tree of the hypotheses. This operator is based on the error covariance generated by the Kalman filter, and has an exponential characteristic. This guarantees that virtually all hypotheses apart from the correct one have a very low interest level and are quickly rejected.

The Kalman filter assumes the target adheres to a motion where acceleration is constant over any three consecutive frames. The Kalman state variable therefore comprises of the centres of mass (CM) of the object over three consecutive frames. The interest operator is defined as

$$I(\Delta_x, \Delta_y) = \mathrm{e}^{-K} \tag{1}$$

where

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

The terms Δ_x, Δ_y are the the differences in the x and y coordinates between the target's CM as predicted by the Kalman tracker and the observed CM of a candidate object. The σ_x, σ_y terms refer to standard deviations over previous predictions. The constant parameter α is a measure of how far away the Δ terms are from the previous observed errors in terms of multiples of standard deviations of those previously observed errors. As α increases the Δ_x, Δ_y errors must be correspondingly smaller for the same level of interest. This is because as α increases, the interest function becomes more peaked. The value $\alpha = 2$ signifies that $I(\Delta_x, \Delta_y)$ gives good response when the inputs are within 0.5 standard deviations from the Kalman covariance.

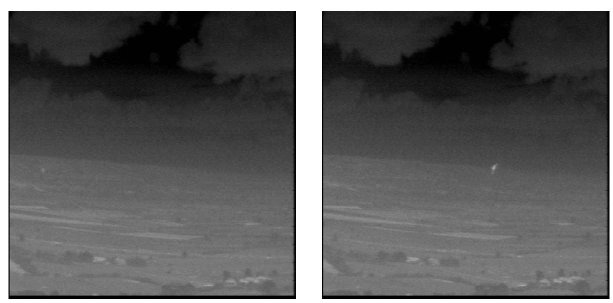
Consequently, the longer the object is tracked well the more robust the interest operator becomes in identifying the true object in the presence of other close and/or high contrast decoys, because the error behaviour itself is modelled more consistently.

3 EXPERIMENTAL RESULTS

Figure 1 shows two frames from our first test sequence of FLIR images in which a target moves from the left side of the scene (starting from one pixel size) to the top right of the scene (ultimately becoming about 40 pixels in size) in a fairly sharp trajectory. Figure 2 shows the detected object and it's tracked trajectory sequence. We detect the target when it is 6-7 pixels in size. Figure 3 shows the profile of the object when it is far away. We emphasise that these images have been considerably enhanced to show the detail. The original FLIR images, as used by the algorithm, had too small a contrast to see the aircraft clearly. The white boxes are the borders of the regions produced by the algorithm for covering the object that is being tracked. The black crosses indicate the positions of the predictions of the centre of mass of the expected region in the next frame as formed by the Kalman filter.

Our wavelet algorithm projects the original image onto its scale space representations (up to three levels). The finest wavelet signal must be used when the object is far away and small. The level may be coarsened as the object becomes larger on approach. Hypotheses at a given resolution level are formed by windows (typically from 4×4 to 8×8 pixels) applied to the relevant wavelet space. Note that no wavelet *synthesis* is performed: we are not concerned with image reconstruction, only in using the wavelet analysis to suggest hypotheses in the form of regions in the original image. A fast connectivity analysis algorithm is then applied to label the connected parts of the chosen region, and these constituted the hypotheses. The interest operator then guides the Kalman tracker into making a decision as to which of the hypotheses is most likely to be the target object.

Typically the Kalman filter produced errors of the order of a few pixels. We show in Figure 4 a plot of errors in the Kalman prediction throughout the sequence.



(a) frame 28 (b) frame 60

Figure 1: Typical frames of a small, long-range target

In Figures 5-9 we show some results from another sequence where the airplane we are tracking encounters another aircraft which could confuse the algorithm. The images show that although many hypothesised targets are generated, the use of our interest operator keeps the algorithm focused on the correct target at all times. For this sequence we used the wavelet coefficients that are

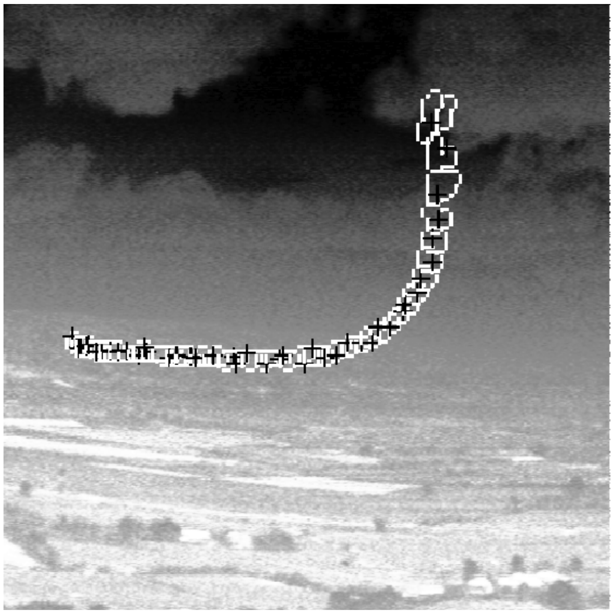


Figure 2: Trajectory of low-contrast target

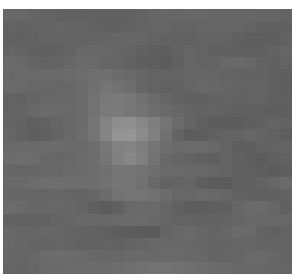


Figure 3: Profile of small object (6-7 pixels) in frame 32

only sensitive to *horizontal* features, because the background was relatively homogeneous and the small plane has greater representation in this space.

4 CONCLUSIONS

Our algorithm combines the discriminating power of the wavelet transform with a Kalman filter to track the motion of small, low contrast objects in image sequences. The method is robust in that the target is always kept in focus, even when higher contrast decoys are present, due to the interest operator. We can also track the object when partial occlusion occurs using the predictions from our adapted Kalman tracker; due to lack of space this issue has not been dealt with here. More details will be available in [5].

The method exploits the simplicity of the fast wavelet transform and the robustness of the Kalman method. We have developed a predictive model that produces reliable estimates of the target's motion.

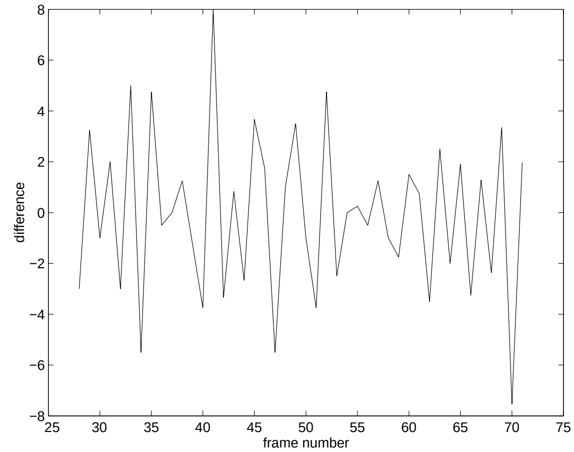


Figure 4: Error between predicted and observed centres of mass of object expressed in the number of pixels

Figure 5: Typical original frame (19)

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 characterization 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 [8] 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 [9].

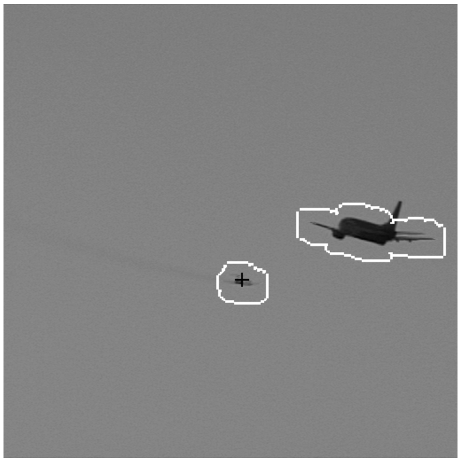


Figure 6: Multiple hypothesis generation (frame 19)

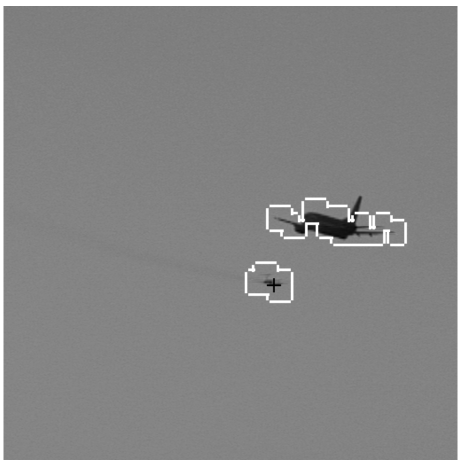


Figure 7: Multiple hypothesis generation (frame 23)

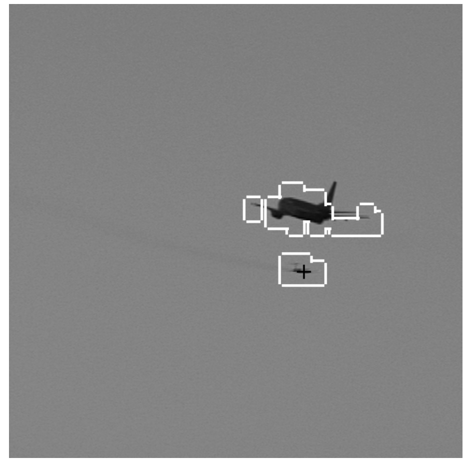


Figure 8: Multiple hypothesis generation (frame 26)

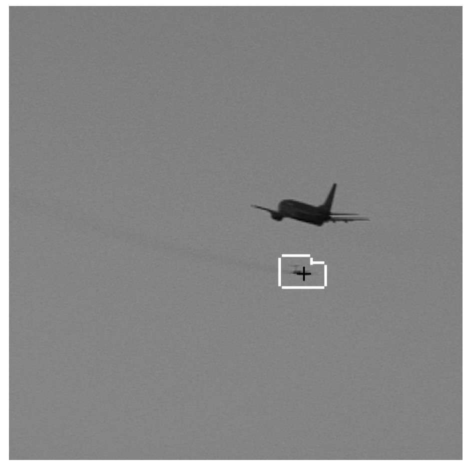


Figure 9: correct tracking of target object

References

[1] D.Casasent & A.Ye. Detection filters and algorithm fusion for ATR. *IEEE Trans. on Image Processing*, 6(1):114–125, 1997.

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

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

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

[5] D. Davies, P.L. Palmer, and M. Mirmehdi. Detection and tracking of very small targets. 1998. *Submitted to the 9th British Machine Vision Conference*, 1998.

[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] S.G. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. on PAMI*, 11(7):674–693, 1989.

[8] 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.

[9] 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.

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