Image-based analysis enables aerial target detection

Enrique Estalayo, Luis Salgado, Fernando Jaureguizar,
and Narciso García

A new strategy reduces computational costs and improves the applicability of automatic aerial target detection.

Detecting and tracking both moving and stationary targets in Forward-Looking Infrared (FLIR) imagery is a challenging research area in computer vision. In contrast to visual images, those obtained from an infrared sensor have extremely low signal to noise ratios (SNR), providing limited information for performing detection or tracking tasks. In addition, while the techniques used to detect moving targets are often based on the static camera hypothesis,¹ sensors used in automatic target recognition applications are typically mounted on moving vehicles such as airplanes, resulting in instabilities during the image acquisition process.¹⁻⁷

The generation of automatic target detection systems has often been application driven and case specific, and has mainly focused on processing terrestrial sequences, with only minor efforts for aerial and maritime environments, as in Meier.⁸ Furthermore, the high complexity of the developed algorithms forces users to accept very high computational costs in order to achieve off-line detection of targets in a scene.

There are many and varied techniques for digitally estimating camera motion and stabilizing the sequence of images, although most of these methods can be classified within two main categories.^{9,10} The first type, flow-based algorithms,⁴⁻⁶ entail very high computational costs, making them impractical for real-time applications. The second type, feature-based methods,³ reduce the computational burden, but do so at the expense of reducing their applicability. Once a sequence is stabilized, there are many different approaches to detect and track the targets present in the scene. However, all of these approaches are still very case specific and have typically been applied only to terrestrial FLIR sequences. To overcome these limitations, we propose an innovative and efficient strategy with a block-based motion estimation and an affine transformation to recover from ego-motion.

Figure 1. The system consists of three major steps: lower resolutions of FLIR images are generated; digital image stabilization is performed; and the detection system is applied to the stabilized images.

This strategy operates on a multi-resolution approach, and improves on the work by Seok,³ minimizing the limitations derived from its oversimplified rotation model. The original model combines rotational and translational motions. The novelty of our strategy relies on the relaxation of the assumed hypothesis, and hence on the enhancement of its applicability, by overcoming the imposition of rotational displacements within the camera motion. In addition, we reduce computational cost by applying a multi-resolution algorithm in which the stabilization technique is applied on the lowest resolution image. After the images have been compensated on the highest resolution level and refined to avoid distortions produced by the sampling process, a dynamic differences-based segmentation is applied, followed by a morphological filtering strategy.

The system described in this paper is composed of three subsystems, as presented in Figure 1. First, a multi-resolution algorithm is applied to the sequence to obtain lower-resolution reproductions of the FLIR images. Next, the digital image stabilization (DIS) system is applied. This consists of two main modules: one for motion estimation and another for motion compensation. The motion estimation module can also be divided: local motion estimation calculates the movement

Continued on next page

Figure 2. The system detects ROIs within the stabilized images (top row) and correctly segments targets within them (bottom row).

of individual image pixels between two consecutive images through a block-matching algorithm; motion type estimation determines whether the displacements correspond to a pure translational motion, a rotational motion, or both at the same time; and the final segment performs global motion estimation. The second DIS module, motion compensation, removes the undesired ego-motion previously estimated. Finally, after compensating in the highest resolution level and refining, the detection system is applied. This is composed of an image differences module, which allows the segmentation of the targets, and a morphological filtering module, which determines their final shape and location within the image.

The fidelity of the image stabilization technique was evaluated using the peak signal-to-noise ratio (PSNR) measure on both synthetically generated and real sequences. Several conclusions can be drawn from the results obtained. First, the motion type estimation module correctly estimates global motion, even in situations with small transformations (e.g. with small rotation angles). Second, results for pure translations and pure rotations are accurate: the estimated values are very similar to those used in the simulations. Finally, the DIS system has demonstrated that our strategy can accurately stabilize the images from real aerial sequences for further processing.

The detection system has been tested on stabilized images from the evaluated sequences. Figure 2 shows some of the results obtained. The detected regions of interest (ROI) containing the potential targets are shown in the top row of images, while the segmented targets are presented in the bottom row. These demonstrate the accuracy of the implemented approach. First, the ROIs were well detected in the sequences, including the

synthetic sequence in which the detection process is more difficult due to the characteristics of the selected image. Second, the targets have been segmented accurately: even in Figure 2(a), which includes extreme contrasts between hot and cold spots. Despite these challenges, the shape of the aircraft was extracted.

These results demonstrate the correct operation of our system, both in stabilizing images and automatically detecting aerial targets from both synthetic and real FLIR sequences. Currently, we are proceeding with further evaluations, specially focused on the generation of a complete automatic target recognition real-time system.

Author Information

Enrique Estalayo, Luis Salgado, Fernando Jaureguizar,
and Narciso García

Image Processing Group

Universidad Politécnica de Madrid

Madrid, Spain

<http://www.gti.ssr.upm.es/~eeg>

<http://www.gti.ssr.upm.es/>

Mr. Enrique Estalayo majored in 2003 in Electrical Engineering at the Universidad Politécnica de Madrid (UPM). For the past two years, he has worked in the Image Processing Group, becoming a PhD Student in the Signals, Systems and Radiocommunications Department. Since November 2005, he has worked for Telefonica R&D, in the development of new digital TV applications.

References

1. S. S. Young, H. Kwon, S. Z. Der, and N. M. Nasrabadi, *Adaptive Target Detection in Forward-Looking Infrared Imagery using the Eigenspace Separation Transform and Principal Component Analysis*, **Opt. Eng.** **43**, pp. 1767–1776, August 2004.
2. A. Yilmaz, K. Shafique, and M. Shah, *Target Tracking in Airborne Forward Looking Infrared Imagery*, **Image and Vision Computing Journal (IVC)** **21** (7), pp. 623–635, 2003.
3. H. D. Seok and J. Lyou, *Digital Image Stabilization using Simple Estimation of the Rotational and Translational Motion*, **Proc. SPIE** **5810**, pp. 170–181, 2005.
4. A. Strehl and J. K. Aggarwal, *MODEEP: a Motion-Based Object Detection and Pose Estimation Method for Airborne FLIR Sequences*, **Machine Vision and Applications** **11** (6), pp. 267–276, 2000.
5. M. Irani, B. Rousso, and S. Peleg, *Recovery of Ego-Motion using Region Alignment*, **IEEE Trans. on Pattern Analysis and Machine Intelligence** **19** (3), pp. 268–272, 1997.
6. J. Y. Chang.
7. C. Marimato and R. Chellappa, *Evaluation of Image Stabilization Algorithms*, **Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing** **5**, pp. 2789–2792, 1998.
8. W. Meier and H. D. vom Stein, *Estimation of Object and Sensor Motion in Infrared Image Sequences*, **Proc. IEEE Int. Conf. on Image Processing** **1**, pp. 568–572, 1994.
9. C. Q. Davis, Z. Z. Karu, and D. M. Freeman, *Equivalence of Subpixel Motion Estimators Based on Optical Flow and Block Matching*, **Proc. IEEE Int. Symp. on Computer Vision**, pp. 7–12, 1995.
10. K. R. Rao and J. J. Hwang, **Techniques and Standards for Image Video and Audio Coding**, Ed. Prentice Hall, PTR, 1996.