

Quantifying Infrared Target Signature Evolution Using AM-FM Features

Mark Shook, John Junger, Nick Mould, Joseph P. Havlicek
School of Electrical and Computer Engineering
University of Oklahoma
Norman, Oklahoma, U.S.A.
baonus@ou.edu, jjunger@ou.edu, stupid@ou.edu, joebob@ou.edu

Abstract—In this paper we combine a new method of measuring infrared target signature evolution with current research and developmental tracking algorithms. Thermal images are decomposed by a set of Gabor filters and demodulated to produce a set of spatio-spectrally localized AM-FM functions corresponding to oriented texture regions from within the original image. Critical updates are detected and issued to a particle filter based tracker, operating *only* on a modulation domain target model, by applying an empirically determined threshold to a new target evolution measurement introduced in this paper. We achieve results comparable to several other theoretical tracking algorithms at a significantly reduced computational cost by eliminating the need to perform parallel tracking in both the pixel and modulation domains.

Keywords-AMCOM, AM-FM Models, Target Tracking

I. INTRODUCTION

We consider the problem of tracking infrared targets immersed in highly structured clutter, where it is common for targets to appear camouflaged by elements of the background that are similar in configuration to those that compose notable features of the target. Under these circumstances, it is difficult to maintain a representation of the target signature that is both uncorrupted by the inclusion of erroneous background features and simultaneously allowed to mature in correspondence with the often profound evolution of the target signature. We continue to use the *AMCOM* library of infrared closure sequences to evaluate the performance of tracking algorithms because they are widely known and provide a large collection of difficult infrared target engagement scenarios. [1]–[4].

Recently, modeling infrared imagery in the modulation domain has been demonstrably verified as a powerful new technique for separating complex target and background structures. Carefully designed filtering and demodulation processes, where the original image is represented by a collection of spatio-spectrally localized modulation functions, have been effectively used to discriminate between infrared targets and backgrounds [5]–[7]. A previous analysis of the *AMCOM* sequences revealed that trackers operating in the pixel and modulation domains generally diverge

differently when the targets evolve beyond the ability of the template to characterize them well. By exploiting the known differences in the failure modes associated with each tracker we have achieved high quality, robust tracking results against a large percentage of the *AMCOM* sequences [8]–[10].

In this paper we present a new method for measuring infrared target signature evolution based on a focused analysis of the texture components isolated in the modulation domain representation. We employ a biologically motivated Gabor filterbank to decompose each pixel domain image acquired by an infrared imaging sensor into multiple oriented narrowband components that are then demodulated to obtain AM and FM functions on a componentwise basis. The modulation domain representation of the target signature is considered in the context of a Bayesian state estimation problem, where a template based auxiliary particle filter is formulated to track the evolution of the target signature over time. Significant variation in the target signature is detected by thresholding our new target signature evolution measurement and used to issue template updates to the tracker.

II. TARGET SIGNATURE REPRESENTATION

Application of AM-FM image models is an established technique for characterizing locally coherent image structure in terms of amplitude (AM) and frequency (FM) modulation functions [11]. Computation of the modulation domain representation for a real-valued image composed of many complex multipartite textures is generally performed by analyzing the image with a filterbank designed to produce locally coherent filter response images, followed by estimation of the AM-FM functions for each of the filterbank channels.

We assume that infrared frames f_k containing a target of interest are captured by the focal plane array of an infrared imaging sensor and delivered to the track processor at time k . The incoming frames are analyzed by a biologically inspired $M = 18$ channel Gabor filterbank similar to those presented in [12] resulting in 18 filter response images y_m , from which the modulating functions are estimated according to

$$\nabla\varphi_m \approx \operatorname{Re} \left[\frac{\nabla(y_m + j\mathcal{H}[y_m])}{j(y_m + j\mathcal{H}[y_m])} \right]$$

This work was supported in part by the U.S. Army Research Laboratory and the U.S. Army Research Office under grant W911NF-08-1-0293.

and

$$a_m \approx \left| \frac{y_m + j\mathcal{H}[y_m]}{G_m[\nabla\varphi_m]} \right|,$$

where $\mathcal{H}[\cdot]$ is the partial Hilbert transform [11].

To represent the state of the target we introduce a joint motion/appearance target model composed of a six element target state vector \mathbf{x}_k and a modulation domain target signature template T . The target state vector \mathbf{x}_k is given by

$$\mathbf{x}_k = [\mathbf{x}_{1,k}^T \ \mathbf{x}_{2,k}^T]^T = [x_{1,k} \ \dot{x}_{1,k} \ \delta_{1,k} \ x_{2,k} \ \dot{x}_{2,k} \ \delta_{2,k}], \quad (1)$$

where $\mathbf{x}_{1,k} = [x_{1,k} \ \dot{x}_{1,k} \ \delta_{1,k}]^T$ and $\mathbf{x}_{2,k} = [x_{2,k} \ \dot{x}_{2,k} \ \delta_{2,k}]^T$. In (1) $x_{1,k}$ and $x_{2,k}$ are the horizontal and vertical coordinates of the target centroid measured in pixels, $\dot{x}_{1,k}$ and $\dot{x}_{2,k}$ are the corresponding horizontal and vertical velocities of the target centroid measured in pixels per frame, and $\delta_{1,k}$ and $\delta_{2,k}$ are the width and height of the target in pixels. The target appearance template T is composed of 18 equally sized rectangular tiles t_m extracted from the 18 a_m images. We define $H(\cdot)$ to be the function that extracts rectangular tiles of raw pixel values from the 18 a_m images at the location and size specified by the state vector \mathbf{x}_k . The modulation domain target signature template T is given by

$$T = \{t_m, m \in [1, M]\} = H(\mathbf{x}_k, f_k),$$

where m indicates channel number.

The initial target state \mathbf{x}_0 is obtained from either an independent detection process or it is manually designated by a human in the loop. In the examples presented here, the initial target state \mathbf{x}_0 is obtained from ground truth data and used to construct the corresponding initial target signature template $T_0 = H(\mathbf{x}_0, f_0)$.

III. TARGET STATE ESTIMATION

We use an auxiliary particle filter similar to those described in [13]–[15] to estimate the target state transition over time. In contrast to SIR (sample importance resampling) particle filters, auxiliary particle filters perform an additional resampling step at each iteration of the filter to eliminate outlying state estimates. The weights used in the additional resampling stage are obtained by prematurely propagating the particles and evaluating their likelihood based on the current observation. Let $\mathbf{v}_k = [v_{1,k} \ u_{1,k} \ 0 \ v_{2,k} \ u_{1,k} \ 0]^T$, $v_{1,k}$, where $v_{2,k}$, $u_{1,k}$ and $u_{2,k}$ are uncorrelated zero-mean Gaussian noise processes and let Δ equal the inter-frame time.

Let $\mathbf{v}_k = [v_{1,k} \ u_{1,k} \ 0 \ v_{2,k} \ u_{1,k} \ 0]^T$, $v_{1,k}$, where $v_{2,k}$, $u_{1,k}$ and $u_{2,k}$ are uncorrelated zero-mean Gaussian noise processes and let Δ equal the inter-frame time. The state transition equation for the target is based on a white noise acceleration model and ternary magnification model given by

$$\begin{bmatrix} \mathbf{x}_{1,k+1} \\ \mathbf{x}_{2,k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{F} & \mathbf{0} \\ \mathbf{0} & \mathbf{F} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{1,k} \\ \mathbf{x}_{2,k} \end{bmatrix} + \mathbf{v}_k, \quad (2)$$

where

$$\mathbf{F} = \begin{bmatrix} 1 & \Delta & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (1 + \gamma) \end{bmatrix} \quad (3)$$

and

$$\gamma = \begin{cases} -\alpha, & p = 1/3, \\ 0, & p = 1/3, \\ \beta, & p = 1/3. \end{cases} \quad (4)$$

In (4) α , 0, and β are the magnification gains with equal probability $p = 1/3$.

The measurement equation is given by $\mathbf{z}_k = \mathbf{H}(\mathbf{x}_k + \mathbf{n}_k, f_k)$, where $\mathbf{n}_k = [n_{1,k} \ n_{2,k}]$, and we assume that $n_{1,k}$ and $n_{2,k}$ are uncorrelated zero-mean Gaussian noise processes. A common modulation domain template $T = \{t_m, m \in [1, M]\}$ is shared by the entire particle population. For each particle \mathbf{x}_k^i , the likelihood $p(\mathbf{x}_k^i | \mathbf{z}_k)$ is determined by computing the normalized cross correlation (NCC) between each plane obtained by resizing T according to the magnification parameters in \mathbf{x}_k^i and the observed tile $H(\mathbf{x}_k^i, f_k)$, where i is the particle index. The likelihood of each particle is determined by averaging the 18 NCC values.

IV. QUANTIFYING TARGET SIGNATURE EVOLUTION

Template based trackers require well planned updates in order to maintain accurate representations of the target signature. In the literature this is known as the template update problem, and many solutions such as updating the template every frame or updating at fixed intervals have been proposed [16]. For many practical tracking sequences, updating the template too often or at poorly chosen times will result in the template becoming corrupted by background clutter which typically causes the tracker to fail.

Here, we present a powerful new metric to quantify temporal evolution of target signatures, where subsequent application of an empirically determined threshold is used to detect major variations in the target signature and issue critical template updates. For each frame k we define the output of the auxiliary particle filter to be the average state vector $\hat{\mathbf{x}}_k$ computed with respect to the population $\hat{\mathbf{x}}_k^i$. We define an 18-element correlation vector $\mathbf{C}_k = NCC[\hat{T}, H(\hat{x}_k, f_k)]$ where \hat{T} is a bicubic resize of the current template T calculated using the magnification parameters of $\hat{\mathbf{x}}_k$.

To characterize acceptable evolution of the target we maintain a running sum of the per channel correlation vectors between the previous template update index j and the current frame k . Evolution of the target signature is detected by thresholding the NCC between the normalized sum of historical template-to-target correlation vectors $\sum_{l=j}^{k-1} \mathbf{C}_l / \|\sum_{l=j}^{k-1} \mathbf{C}_l\|$ and the correlation metric from the current frame \mathbf{C}_k .

Table I
ABSOLUTE ERROR IN TRACKED CENTROID LOCATION (MEASURED IN PIXELS) AND NUMBER OF TEMPLATE UPDATES PERFORMED.

Sequence	Number of Frames	Error	Number of Updates
<i>mg14_15</i>	135	2.718	4
<i>mg16_18</i>	135	2.194	4
<i>mg18_16</i>	179	3.764	15
<i>mg19_06</i>	165	2.847	16

V. EXPERIMENTAL RESULTS

We performed auxiliary particle filtering on modulation domain target models using the new template update strategy against four of the most challenging *AMCOM* sequences. Table I shows the error in the tracked centroid in pixels and the number of template updates performed during the sequence. The first column in Fig. (1) shows the track gate specified by \hat{x}_k at the times when template updates were performed. The second and third columns show pixel domain tiles that correspond to the target signature template being used by the tracker and the current target signature from the perspective of the tracking algorithm, respectively. The fourth column of Fig. (1) is the squared error between the current template (column 2) and the tracked target (column 3), computed for the sake of highlighting the visible structural differences between the template and the target. Although the tracking algorithm processes 18 plane modulation domain imagery, we have chosen to display pixel domain targets and images to save space.

In this paper we have introduced powerful new target representation and analysis techniques for quantifying infrared target signature evolution over time. Computation of AM-FM image models for multiple spatio-spectrally localized subband images derived by analysis of infrared imagery produces a collection of AM-FM features that correspond well to the oriented texture regions within the pixel domain image acquired by the infrared sensor. A Comparison between current and historical template-to-target correlation structures is used to identify and issue meaningful updates to the track processor resulting in excellent tracking performance on several challenging *AMCOM* sequences.

REFERENCES

- [1] J.F. Khan and M.S. Alam, "Efficient target detection in cluttered FLIR imagery," in *Optical Pattern Recognition XVI*, D.P. Casasent and T.-H. Chao, Eds., 2005, vol. 5816 of *Proc. SPIE*, pp. 39–53.
- [2] A. Dawoud, M.S. Alam, A. Bai, and C. Loo, "Decision fusion algorithm for target tracking in infrared imagery," *Opt. Eng.*, vol. 44, no. 2, 2005.
- [3] S. Yi and L. Zhang, "A novel multiple tracking system for UAV platforms," in *ISR Systems and Applications III*, D.J. Henry, Ed., 2006, vol. 6209 of *Proc. SPIE*.
- [4] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *ACM Comput. Surv.*, vol. 38, no. 4, pp. 1–45, Dec. 2006.
- [5] C.T. Nguyen and J.P. Havlicek, "Modulation domain features for discriminating infrared targets and backgrounds," in *Proc. IEEE Int'l. Conf. Image Proc.*, Atlanta, GA, Oct. 8–11, 2006, pp. 3245–3248.
- [6] J.P. Havlicek, C.T. Nguyen, and M. Yeary, "Modulation domain infrared target models," in *Targets and Backgrounds XII: Characterization and Representation*, W. R. Watkins and D. Clement, Eds., 2006, vol. 6239 of *Proc. SPIE*, pp. 62390D–1 – 62390D–11.
- [7] R. S. Prakash and R. Aravind, "Modulation-domain particle filter for template tracking," in *Proc. Int'l. Conf. Pattern Recog.*, Tampa, FL, Dec. 8–11, 2008, pp. 1–4.
- [8] N.A. Mould, C.T. Nguyen, C.M. Johnston, and J.P. Havlicek, "Online consistency checking for AM-FM target tracks," in *Proc. SPIE/IS&T Conf. Computational Imaging VI*, C.A. Bouman, E.L. Miller, and I. Pollak, Eds., 2008, vol. 6814 of *Proc. SPIE*, pp. 681413–1 – 681413–12.
- [9] C.M. Johnston, N.A. Mould, and J.P. Havlicek, "Multichannel dual domain infrared target tracking for highly evolutionary target signatures," in *Proc. IEEE Int'l. Conf. Image Proc.*, Cairo, Egypt, 2009.
- [10] N.A. Mould, C.T. Nguyen, and J.P. Havlicek, "Infrared target tracking with AM-FM consistency checks," in *Proc. IEEE Southwest Symp. Image Anal., Interp.*, Santa Fe, NM, Mar. 24–26, 2008, pp. 5–8.
- [11] J.P. Havlicek, D.S. Harding, and A.C. Bovik, "Multidimensional quasi-eigenfunction approximations and multicomponent AM-FM models," *IEEE Trans. Image Proc.*, vol. 9, no. 2, pp. 227–242, Feb. 2000.
- [12] J.G. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *J. Opt. Soc. Am. A*, vol. 2, no. 7, pp. 1160–1169, Jul. 1985.
- [13] M.K. Pitt and N. Shepard, "Filtering via simulation: Auxiliary particle filters," *Journal of the American Statistical Association*, vol. 94, no. 446, pp. 590–599, June 1999.
- [14] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online non-linear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, 2002.
- [15] M.G.S. Bruno, "Bayesian methods for multispectral target tracking in image sequences," *IEEE Trans. Image Proc.*, vol. 52, no. 7, pp. 1848–1861, Jul 2004.
- [16] I. Matthews, T. Ishikawa, and S. Baker, "The template update problem," *IEEE Trans. Pattern Anal., Machine Intel.*, vol. 26, no. 6, pp. 810–815, Jun. 2004.

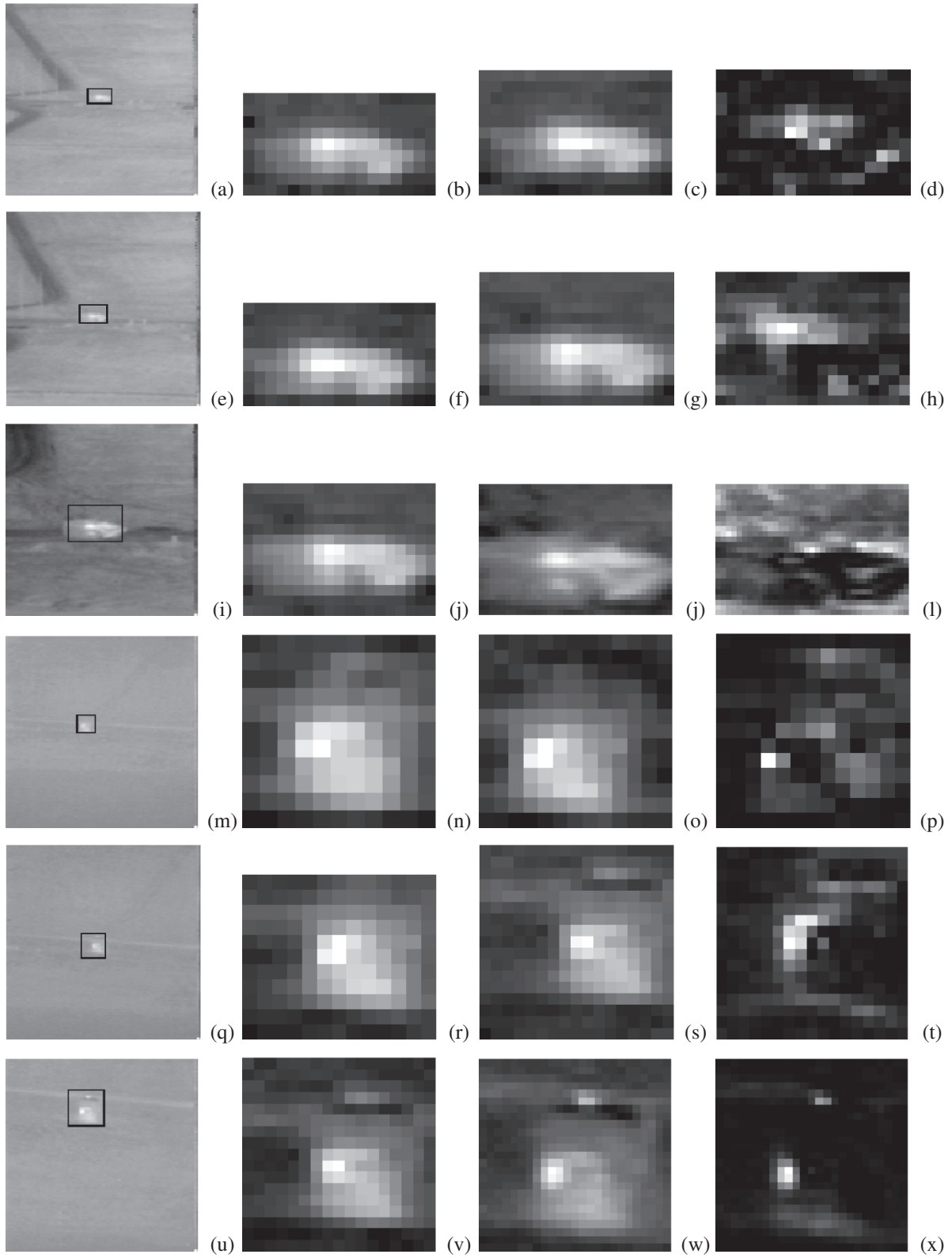


Figure 1. Selected frames from *AMCOM* sequences *rng16_18* (rows 1-3 correspond to frames 140, 151 and 179) and *rng14_15* (rows 4-6 correspond to frames 140, 147 and 177), depicting the tracked target (column 1), the current template used by the tracker (column 2), the current target from the perspective of the tracker (column 3) and the squared error measured between the template and the current target (column 4).