

# Dual Domain Auxiliary Particle Filter with Integrated Target Signature Update

Colin M. Johnston, Nick Mould, and Joseph P. Havlicek  
School of Electrical & Computer Engineering  
University of Oklahoma, Norman, OK 73019, USA

Guoliang Fan  
School of Electrical & Computer Engineering  
Oklahoma State University, Stillwater, OK 74078, USA

## Abstract

*For the first time, we formulate an auxiliary particle filter jointly in the pixel domain and modulation domain for tracking infrared targets. This dual domain approach provides an information rich image representation comprising the pixel domain frames acquired directly from an imaging infrared sensor as well as 18 amplitude modulation functions obtained through a multicomponent AM-FM image analysis. The new dual domain auxiliary particle filter successfully tracks all of the difficult targets in the well-known AMCOM closure sequences in terms of both centroid location and target magnification. In addition, we incorporate the template update procedure into the particle filter formulation to extend previously studied dual domain track consistency checking mechanism far beyond the normalized cross correlation (NCC) trackers of the past by explicitly quantifying the differences in target signature evolution between the modulation and pixel domains. Experimental results indicate that the dual domain auxiliary particle filter with integrated target signature update provides a significant performance advantage relative to several recent competing algorithms.*

## 1. Introduction

Tracking infrared targets in the midwave infrared (MWIR;  $3\mu\text{m}$  -  $5\mu\text{m}$ ) and longwave infrared (LWIR;  $8\mu\text{m}$  -  $12\mu\text{m}$ ) bands is a difficult problem exacerbated by a myriad of factors characteristic of the modern battlespace. Analysis of infrared targets from the well-known AMCOM sequences [8, 4, 16, 17] reveal several specific challenges associated with tracking ground based targets, including: significant sensor platform egomotion, radical target signature

evolution, and significant, unpredictable target kinematics. Accurate infrared target tracking is critical in many military weapons systems where common knowledge indicates that improving infrared target detection and tracking has the potential to simultaneously minimize unwanted collateral damage and maximize the probability of successful target elimination.

The AMCOM closure sequences provide samples of several practical battlefield scenarios involving terrestrial military vehicles as well as a few stationary infrared sources in both the MWIR and LWIR bands. In each AMCOM sequence, an airborne infrared sensor rapidly closes on one or more ground based target(s), thereby producing a sequence of frames exhibiting significant target magnification and egomotion, as well as profound nonstationary target signature evolution in many cases. In addition, the backgrounds in the AMCOM sequences generally contain highly structured clutter that complicates the problem of discriminating between the targets of interest and the respective ambient surroundings in which they are immersed.

Recent approaches to tracking infrared targets amidst highly structured clutter, where the target signature undergoes significant nonstationary evolution, almost always depend on a target signature appearance model (*e.g.*, template) that is used to represent the current appearance of the target of interest in a state space formulation. Template tracking in the modulation domain has also been demonstrated recently [11, 15]. In practice, the target detection processes is often coupled with a tracking algorithm consistent with either a Kalman or Particle filtering methodology [2, 3, 17]. Many such approaches rely on static *a priori* or empirical appearance models and are related to classical matched filtering techniques in that the likelihood of a specific target instance is quantified by the mean squared error (MSE) or the normalized cross correlation (NCC) between the appearance model and the acquired infrared video frames. Inevitably, tracking algorithms that depend on a

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.

target template fail if the target signature that is actually acquired from the sensor evolves to a point where the stored appearance model no longer accurately represents a significant fraction of structure in the observed signature. The problem of refreshing the target template is well known and is referred to in the literature as the *template update problem* [5, 9, 10, 13]. Several strategies for preventing stale templates include updating the template every frame, updating the template at a fixed or variable time interval, and, more recently, updating the template based on track consistency checks which detect divergence of the tracked target centroid in the pixel and modulation domains [11, 12].

The dual domain tracking approach that we introduce in this paper uses a new modulation domain auxiliary particle filter to maintain tight track locks against the most challenging targets in the AMCOM closure sequence data set. The new dual domain track processor maintains three independent particle populations, two of which are used to detect when the target template has become stale and one which tracks the target jointly in the fused modulation and pixel domain feature space. In addition, we improve upon the recent dual domain track consistency checks used in [11, 12], which rely on the distance between the pixel domain and modulation domain target centroids alone, by quantifying the difference between the target signature evolution in the two domains more comprehensively. Although we continue to threshold the distance between the modulation domain and pixel domain centroids, the auxiliary particle filter given here explicitly integrates the observed target appearance into the resampling step that is performed just prior to estimating the target size and location.

## 2. Target Signature Model

We assume a practical target tracking scenario where a midwave or longwave infrared sensor produces a sequence of video frames  $f_k$  at time instants  $k \in \mathbb{Z}$ . In addition, we assume that the initial target designation is obtained from an independent detection process based on a stored signature library or from a human operator. In either case, the location of the target centroid in image coordinates and the horizontal and vertical sizes of the target at time  $k = 0$  are known and are used to initialize the track processor.

The initial pixel domain target template is obtained by extracting a window of pixel values of the given size from about the given centroid in the frame  $f_0$ . Similarly, the modulation domain target template is obtained by extracting amplitude modulation (AM) values from an identical neighborhood of pixel sites in the multicomponent AM-FM model  $\ell_0$  computed from the pixel domain frame  $f_0$  as described below.

To obtain the modulation domain model  $\ell_k$  for any pixel domain frame, we model the frame as a sum of AM-FM components according to  $f_k = \sum_m t_{k,m}$  as in [7]. We

then apply an  $M = 18$  channel Gabor filterbank to  $f_k$  as in [6, 11, 12] to separate 18 components  $t_k$  important to the infrared tracking problem from one another in space and in frequency. This produces 18 complex-valued response images  $y_{k,m} = f_k * g_m$ , where  $g_m$  and  $G_m$  are the impulse response and frequency response of filterbank channel  $m$  and where a frequency domain implementation of the discrete partial Hilbert transform  $\mathcal{H}$  is folded into the implementation of  $G_m$  to efficiently generate the complex-valued responses as described in [7] according to

$$y_{k,m} = (f_k * g_m) + j\mathcal{H}[f_k * g_m] = (f_k + j\mathcal{H}[f_k]) * g_m. \quad (1)$$

Let  $\mathbf{x} \in \mathbb{R}^2$ . As in [7], the AM functions  $a_{k,m}$  and FM functions  $\nabla\varphi_{k,m}$  of the image components  $t_{k,m}$  may then be estimated from the channel responses  $y_{k,m}$  according to

$$\nabla\varphi_{k,m}(\mathbf{x}) \approx \text{Re} \left[ \frac{\nabla y_{k,m}(\mathbf{x})}{j y_{k,m}(\mathbf{x})} \right], \quad (2)$$

$$a_{k,m}(\mathbf{x}) \approx |y_{k,m}(\mathbf{x}) / G_{k,m}[\nabla\varphi_{k,m}(\mathbf{x})]|. \quad (3)$$

The multicomponent modulation domain model

$$\ell_k = \{a_{k,m}, \nabla\varphi_{k,m}\}_{1 \leq m \leq M}. \quad (4)$$

is then given by the collection of estimated AM and FM functions for the  $M = 18$  components, as depicted schematically in Fig. 1.

When a new target track is initialized at  $k = 0$ , the dual domain target model comprises the pixel domain target template extracted from  $f_0$  and 18 AM function templates extracted from  $\ell_0$ . Although the FM functions in  $\ell_0$  also contain significant information about the local texture structure and were used for target tracking in [11, 15], we have found that the FM functions are relatively more sensitive to target signature evolution making it necessary to adaptively weight the relative contributions of the AM and FM functions in order to maintain robust track locks against the AMCOM sequences. While relative weighting of the AM and FM features was considered in [15] for tracking at visible wavelengths, it is beyond the scope of this paper to propose an analogous weighting scheme capable of tracking the most challenging AMCOM targets. Therefore, only the AM functions are included in our dual domain feature space here. Note that the FM functions are not superfluous, however, and contribute to the estimation of the AM functions through (3).

## 3. Tracking Algorithm

We apply an auxiliary particle filter in the modulation domain for the first time. Compared to the importance resampling (SIR) particle filter, the auxiliary particle filter produces a population of particles that are grouped more

tightly about the state vector sample mean by removing inefficient outliers in the second stage weight assignment process [1, 3, 14]. The auxiliary particle filter proposed here is initialized with three independent populations of particles which we refer to as the *pixel domain particles*, the *modulation domain particles*, and the *dual domain particles*. Weights for the pixel domain particles are updated using only pixel domain observations, whereas those for the modulation domain particles are updated using only modulation domain observations and those for the dual domain particles are updated using both pixel domain and modulation domain observations. The state vectors of the particles in all three populations contain position variables that are initially Gaussian distributed about the initial target centroid designation.

The dual domain track processor maintains a dual domain target template that is obtained by stacking together the pixel domain template and the 18 modulation domain AM templates. Subsequent to the template initialization in frame  $f_0$ , all 18 planes of this dual domain template are refreshed together in any frame where the expected target centroid computed with respect to the pixel domain particle population diverges from that computed with respect to the modulation domain particle population by more than three pixels. To ensure statistical independence among the three particle populations resampling and weighting is performed independently on each population.

### 3.1. State Model

Target kinematics and magnification for the auxiliary particle filter are modeled by a six component state vector

$$\mathbf{x}_k = [x_{1,k} \ \dot{x}_{1,k} \ \delta_{1,k} \ x_{2,k} \ \dot{x}_{2,k} \ \delta_{2,k}]^T = [\mathbf{x}_{1,k}^T \ \mathbf{x}_{2,k}^T]^T, \quad (5)$$

where  $x_{1,k}$ ,  $\dot{x}_{1,k}$ , and  $\delta_{1,k}$  are the horizontal position, velocity, and magnification of the target,  $x_{2,k}$ ,  $\dot{x}_{2,k}$ , and  $\delta_{2,k}$  are the vertical position, velocity, and magnification, and where  $\mathbf{x}_{1,k} = [x_{1,k} \ \dot{x}_{1,k} \ \delta_{1,k}]^T$  and  $\mathbf{x}_{2,k} = [x_{2,k} \ \dot{x}_{2,k} \ \delta_{2,k}]^T$ . The magnification parameters  $\delta_{1,k}$  and  $\delta_{2,k}$  give the target size hypothesized by each individual particle relative to the size of the most recently updated global target template, *i.e.*, the magnification parameters are used for scaling the global template to the appropriate width and height.

A constant velocity model is coupled with a uniform ternary process to model the motion of the target centroid and the observed magnification. The state transition equation used to propagate the particles in all three particle populations is given by

$$\begin{bmatrix} \mathbf{x}_{1,k+1} \\ \mathbf{x}_{2,k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{F} & \mathbf{0} \\ \mathbf{0} & \mathbf{F} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{1,k} \\ \mathbf{x}_{2,k} \end{bmatrix} + \mathbf{v}_k, \quad (6)$$

where

$$\mathbf{F} = \begin{bmatrix} 1 & \Delta & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (1 + \gamma) \end{bmatrix}, \quad (7)$$

The noise vector  $\mathbf{v}_k = [v_{1,k} \ 0 \ 0 \ v_{2,k} \ 0 \ 0]^T$  is made up of uncorrelated zero mean white Gaussian noise variables  $v_{1,k}$  and  $v_{2,k}$  and the uniform ternary process  $\gamma$  that governs magnification of the target appearance model is defined by

$$\gamma = \begin{cases} -\alpha, & p = 1/3, \\ 0, & p = 1/3, \\ \beta, & p = 1/3. \end{cases} \quad (8)$$

where  $\alpha$  and  $\beta$  are magnification gain parameters. In general we set  $\alpha = \beta = 0.10$ . However, for certain AMCOM closure sequences such as *rng14.15* improved tracking performance is obtained by choosing  $\alpha = 0.05$  and  $\beta = 0.10$ .

The observation equation is given by  $\mathbf{z}_k = G(\mathbf{x}_k, \mathbf{n}_k)$ , where  $G(\cdot)$  transforms the target state vector  $\mathbf{x}_k$  into an image by extracting pixels from the current frame at the location and size specified in  $\mathbf{x}_k$ .

### 3.2. Likelihood Function

At each time step  $k$ , each particle from each particle population is resampled and weighted according to a likelihood function  $p(z_k, x_k^i)$  that indicates the likelihood that a specific particle represents the true target signature well. We define the likelihood by

$$p(z_k, x_k^i) = \exp[-\text{MSE}(I_k^i, \hat{I}_k^i)], \quad (9)$$

where “MSE” indicates the mean squared error between  $I_k^i$  and  $\hat{I}_k^i$  and where  $I_k^i$  is a subset of the global target template resized via bicubic interpolation to the horizontal and vertical magnifications given by the state vector of particle  $x_k^i$  and including only the collection of modulation domain and/or pixel domain planes which are appropriate for the population to which the particle  $x_k^i$  belongs. The observation  $\hat{I}_k^i$  in (9) has the same size as  $I_k^i$  and is extracted from a neighborhood in  $f_k$  and  $\ell_k$  located about the target centroid hypothesized by particle  $x_k^i$  using the same planes that were used to generate  $I_k^i$ .

After weighting each particle according to (9), resampling is performed and the expected state vector for each population is computed by averaging across the particles in that population. Finally, a second weight assignment is performed within each population using the modified likelihood function

$$p(z_k, x_k^i) = \exp[-\text{MSE}(E_k, \hat{I}_k^i)]. \quad (10)$$

Here,  $\hat{I}_k^i$  is the same as in (9), while  $E_k$  is generated from the global target template similarly to  $I_k^i$  in (9), but using the centroid and magnification parameters from the averaged state vector from the population to which  $x_k^i$  belongs. This second weight assignment procedure reduces outliers in the particle populations by re-evaluating the particles based on their (previously unknown) respective average state vectors, thereby reducing the variance of each population.

### 3.3. Template Update

As the observed target signature evolves, the template tends to become stale in the sense that it no longer provides an accurate appearance model for the target. To combat this problem, we quantify the difference in target signatures between the modulation and pixel domains as the difference in centroid locations between the expected values of particle populations in the pixel and modulation domains. If the target signature evolution in the pixel and modulation domains is not consistent, as determined by thresholding the distance between their centroids, then the global template is refreshed by extracting a new template from all 19 planes of  $f_k$  and  $\ell_k$ . This new template is extracted from a neighborhood about the centroid of the average state vector computed with respect to the dual domain particle population and has spatial size given by the magnification parameters of the dual domain average state vector.

## 4. Experimental Results

In this section, we compare the tracking performance of the new dual domain auxiliary particle filter (DDAPF) to that obtained with a pixel domain SIR filter (PDSIR), a modulation domain SIR filter (MDSIR), and a dual domain SIR filter (DDSIR) [11, 12, 15]. First we selected the AMCOM sequences that are most difficult in terms of magnification and rapid target signature evolution in our opinion. The names of these sequences are given in the first column of Table 1. Since all of the AMCOM sequences contain blocks of frames that do not exhibit significant target signature evolution, we specifically selected sequences with long runs of hard to track frames and then removed frames from the beginning and end of each sequences to maximize the percentage frames containing radical target signature evolution. This also minimizes the influence of stagnant target and background frames on the measurement of the tracking algorithm's performance, effectively subjecting all of the tracking algorithms to the severest scenarios available in the AMCOM closure sequences.

At the beginning of each sequence, the target size and location were manually designated based on ground truth. Each of the selected sequences was then processed using the four particle filtering methods described above.

Results from these four track processors are summarized in Table 1, where percentage improvement is calculated using the pixel domain SIR filter results as a baseline. For all of the AMCOM sequences tested the dual domain auxiliary particle filter provided significant performance gains relative to the three SIR filters in terms of the average absolute error in the tracked centroid. Incorporation of the auxiliary particle filter dramatically improved the accuracy of the estimated target size in all sequences tested relative to the competing SIR filters, although reliable quantitative ground

Sequence	DDAPF	DDSIR	MDSIR	PDSIR
<i>rng14_15</i>	<b>2.56</b>	2.86	3.51	3.65
<i>rng15_20</i>	<b>1.26</b>	1.82	2.09	2.03
<i>rng16_18</i>	<b>2.01</b>	2.08	2.15	2.14
<i>rng18_16</i>	<b>1.20</b>	1.23	1.58	1.82
<i>rng19_06</i>	<b>1.27</b>	2.65	2.76	2.89
<i>rng19_07</i>	<b>2.01</b>	2.55	2.83	3.04
<i>rng19_13</i>	<b>2.33</b>	4.39	4.70	4.72
<i>rng19_NS</i>	<b>2.07</b>	2.28	2.85	2.93
% Improve	<b>34.27%</b>	14.77 %	3.12%	0%

Table 1. Average absolute error in tracked centroid position, measured in pixels, for a variety of particle filters run against the most difficult AMCOM sequences in terms of magnification and general target signature evolution. PSIR - Pixel Domain SIR Filter; MDSIR - Modulation Domain SIR Filter; DDSIR - Dual Domain SIR Filter; DDAPF - Dual Domain Auxiliary Particle Filter.

truth data for the target size is not available. For this reason, we do not give a table summarizing the tracked target size versus ground truth. Instead, representative tracked frames from two sequences that are dominated by magnification change are given in Fig. 2, where the track gate estimated by the dual domain auxiliary particle filter is shown overlaid on the raw infrared frames. Eight frames are shown for each of the two sequences *rng14\_15* and *rng16\_18*.

## 5. Conclusion

The new dual domain auxiliary particle filter introduced here for the first time has shown excellent performance gains over pixel domain, modulation domain, and dual domain SIR track filters against several of the most difficult from among the well-known AMCOM closure sequences. The dual domain auxiliary particle filter significantly improved the tracking performance in all of the sequences tested. The data in this table show that the dual domain track processors provide a substantial advantage over the single domain trackers and that the auxiliary particle filter delivers a substantial improvement over any SIR filter.

The dual domain track processor succeeds in fusing pixel domain and modulation domain measurements in a meaningful way to deliver a previously unavailable characterization of infrared targets and backgrounds that is both dense and information rich, facilitating the detection and accurate tracking of profoundly evolutionary signatures and the implementation of practical dual domain track consistency checks. Our future work in this area will include improving the accuracy of the computed AM and FM functions by replacing the Gabor filterbank used here with a steerable pyramid decomposition that supports error- and approximation-free continuous demodulation algorithms based on tensor product splines and the development of new techniques for

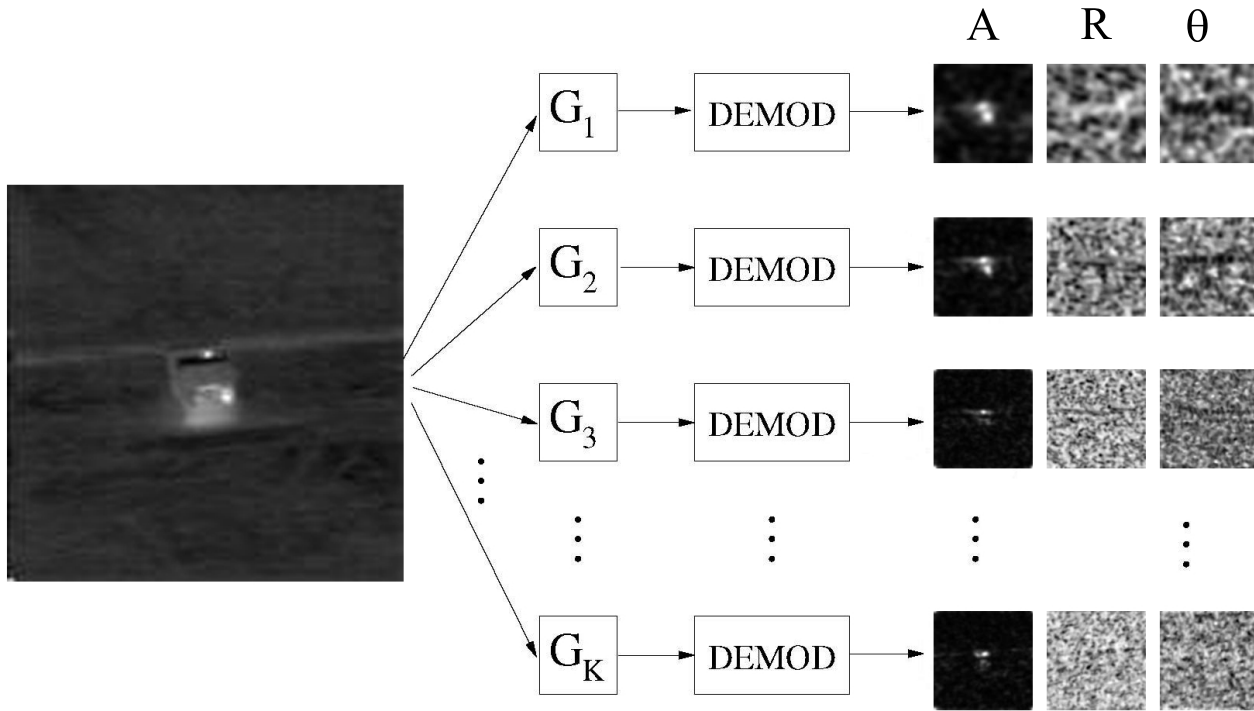


Figure 1. Example of the Gabor analysis filterbank and subsequent demodulation performed on frame 240 of AMCOM sequence *rng14\_15*. The frequency modulating functions are calculated as the derivative of the phase in both the vertical and horizontal directions. In this figure the horizontal and vertical components of the frequency vectors have been transformed into their polar equivalents  $R$  and  $\theta$ .

incorporating the improved FM functions into the dual domain auxiliary particle filtering framework introduced in this paper.

## References

- [1] 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.*, 50(2):174–188, 2002.
- [2] M. Bruno. Sequential importance sampling filtering for target tracking in image sequences. *IEEE Signal Process. Let.*, 10(8):246–249, Aug 2003.
- [3] M. Bruno. Bayesian methods for multispectral target tracking in image sequences. *IEEE Trans. Image Proc.*, 52(7):1848–1861, Jul 2004.
- [4] A. Dawoud, M. Alam, A. Bai, and C. Loo. Decision fusion algorithm for target tracking in infrared imagery. *Opt. Eng.*, 44(2), 2005.
- [5] G. Hager and P. Belhumeur. Real-time tracking of image regions with changes in geometry and illumination. In *Proc. IEEE Int'l. Conf. Computer Vision, Pattern Recog.*, pages 403–410, San Francisco, CA, Jun. 18-20 1996.
- [6] J. Havlicek, A. Bovik, and D. Chen. AM-FM image modeling and Gabor analysis. In C. Chen and Y. Zhang, editors, *Visual Information Representation, Communication, and Image Processing*, pages 343–385. Marcel Dekker, New York, 1999.
- [7] J. Havlicek, D. Harding, and A. Bovik. Multidimensional quasi-eigenfunction approximations and multicomponent AM-FM models. *IEEE Trans. Image Proc.*, 9(2):227–242, Feb. 2000.
- [8] J. Khan and M. Alam. Efficient target detection in cluttered FLIR imagery. In D. Casasent and T.-H. Chao, editors, *Optical Pattern Recognition XVI*, volume 5816 of *Proc. SPIE*, pages 39–53, 2005.
- [9] L. Latecki and R. Miezianko. Object tracking with dynamic template update and occlusion detection. In *Proc. IEEE Int'l. Conf. Pattern Recog.*, pages 556–560, Hong Kong, China, Aug. 20-24, 2006, vol. 1.
- [10] I. Matthews, T. Ishikawa, and S. Baker. The template update problem. *IEEE Trans. Pattern Anal., Machine Intel.*, 26(6):810–815, Jun. 2004.
- [11] N. Mould, C. Nguyen, and J. Havlicek. Infrared target tracking with AM-FM consistency checks. In *Proc. IEEE South-*

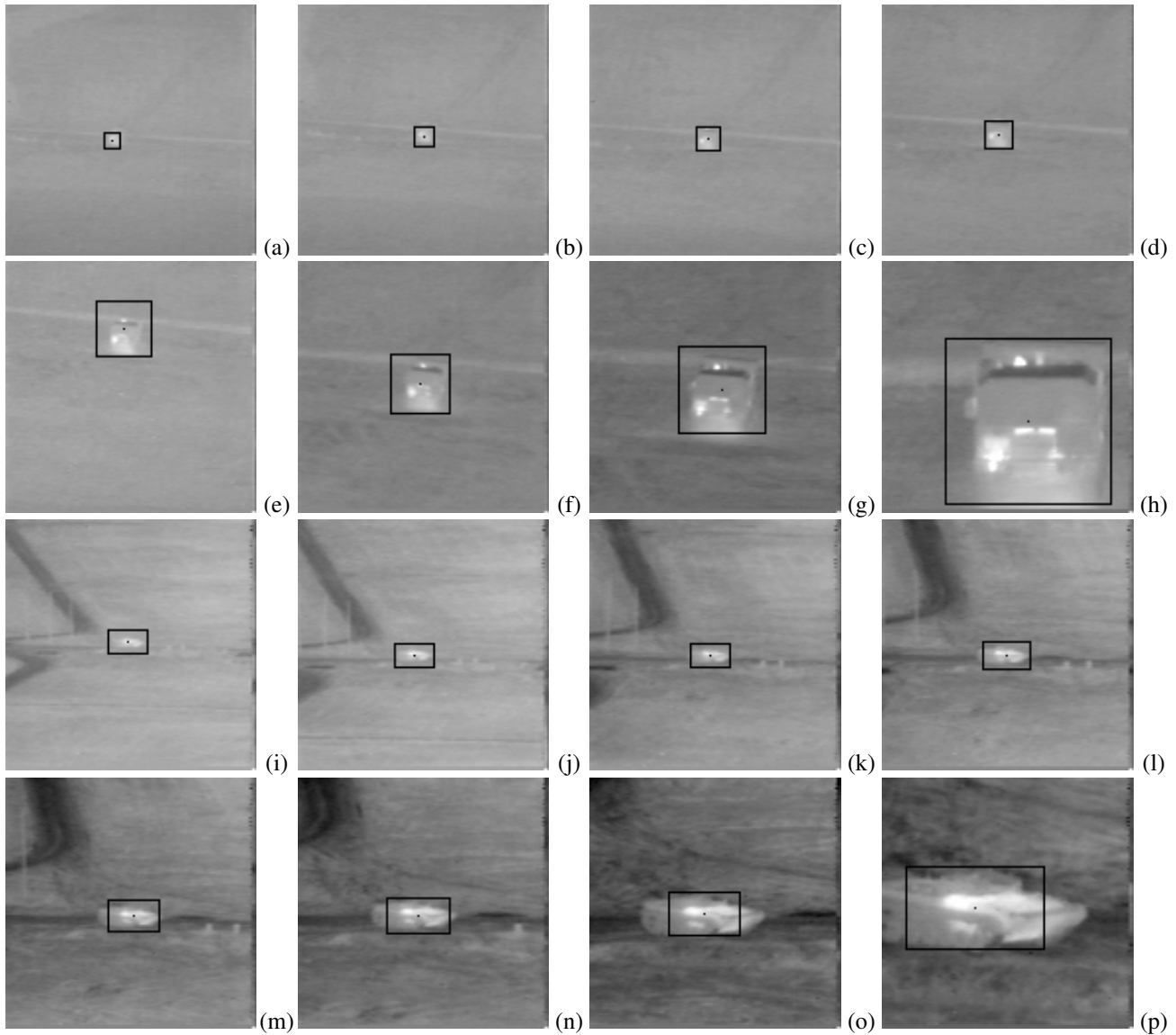


Figure 2. Experimental tracking results using the new dual domain auxiliary particle filter on two difficult AMCOM sequences. *mg14\_15* frames (a) 140, (b) 160, (c) 180, (d) 200, (e) 220, (f) 240, (g) 260, (h) 280 and *mg16\_18* frames (i) 150, (j) 170, (k) 190, (l) 210, (m) 230, (n) 250, (o) 270, and (p) 290.

- west Symp. Image Anal., Interp.*, pages 5–8, Santa Fe, NM, Mar. 24–26, 2008.
- [12] N. Mould, C. Nguyen, C. Johnston, and J. Havlicek. Online consistency checking for AM-FM target tracks. In C. Bouman, E. Miller, and I. Pollak, editors, *Proc. SPIE/IS&T Conf. Computational Imaging VI*, volume 6814 of *Proc. SPIE*, pages 681413–1 – 681413–12, 2008.
- [13] Z. Peng, Q. Zhang, and A. Guan. Extended target tracking using projection curves and matching pel count. *Optical Eng.*, 46(6):066401–1 – 066401–6, Jun. 2007.
- [14] M. K. Pitt and N. Shephard. Filtering via simulation: auxiliary particle filters. *J. Amer. Statistic. Assoc.*, 94(446):590–599, 1999.
- [15] R. S. Prakash and R. Aravind. Modulation-domain particle filter for template tracking. In *Proc. Int’l. Conf. Pattern Recog.*, pages 1–4, Tampa, FL, Dec. 8–11, 2008.
- [16] S. Yi and L. Zhang. A novel multiple tracking system for UAV platforms. In D. Henry, editor, *ISR Systems and Applications III*, volume 6209 of *Proc. SPIE*, 2006.
- [17] A. Yilmaz, O. Javed, and M. Shah. Object tracking: A survey. *ACM Comput. Surv.*, 38(4):1–45, Dec. 2006.