Modulation domain infrared target models

Joseph P. Havlicek, Chuong T. Nguyen, and Mark Yeary

University of Oklahoma, Norman, OK, USA

ABSTRACT

We compute joint AM-FM models that characterize infrared targets and backgrounds in the modulation domain. We consider spatially localized structures within an IR image as sums of nonstationary, quasi-sinusoidal functions admitting locally narrowband amplitude and frequency modulations. By quantitatively estimating the modulations that dominate the signal spectrum on a spatially local basis, we obtain a new modulation domain feature vector that can augment the more traditional pixel domain, Fourier spectrum, and multispectral color features that have been used in IR target detection and tracking systems for a long time. Our preliminary studies, based primarily on midwave and longwave missile approach sequences, suggest that IR targets and backgrounds do typically possess sufficient spatially local modulated structure (i.e., texture) for modulation domain techniques to be meaningfully applied. We also present qualitative results strongly indicating that the modulation domain feature vector is a powerful tool for discriminating infrared targets and backgrounds.

Keywords: AM-FM models, filter banks, Gabor filters, modulation domain, target discrimination

1. INTRODUCTION

The notions of amplitude and frequency modulation are fundamental in telecommunications and speech processing, where one normally thinks of a high frequency carrier signal that is modulated by one or more relatively lower frequency information signals. While the dichotomy between carrier and information signal does not apply as readily to typical visible images, it is well-known that texture may be interpreted as a carrier of region information in the classical image segmentation problem. AM-FM image models, which consider an image as a sum of nonstationary quasi-sinusoidal components, have been used with great success in a variety of image processing and computer vision applications including, e.g., texture segmentation, shape from texture, content based image retrieval, and stereopsis. Computed estimates of the amplitude and frequency modulating functions provide a powerful description of the structure of an image in the modulation domain. In this paper, we apply AM-FM modeling techniques to compute modulation domain descriptions of infrared targets and backgrounds for the first time. Our goals are to determine if typical infrared imagery possesses sufficient texture structure to be meaningfully treated by modulation domain techniques and to provide an assessment of the potential of modulation domain feature vectors for enhancing the performance of traditional target detection, tracking, and identification processes. In Section 2, we motivate the notions of instantaneous amplitude, phase, and frequency, and provide a brief overview of AM-FM image modeling techniques. In Section 3, we consider the problem of computing AM-FM models for infrared imagery and develop adaptive Gabor filters for extracting estimates of the dominant modulations of infrared targets. Two comprehensive examples are given in Section 4, where we compute modulation domain feature vectors from longwave infrared missile closure sequences.

2. AM-FM IMAGE MODELS

The classical Fourier transform represents a signal as a sum of complex sinusoids, each one of which has a constant amplitude and a constant frequency. Thus, a signal $x : \mathbb{R} \to \mathbb{C}$ may be written in terms of the uncountable set of spectral basis signals $\{e^{j\omega t}\}_{\omega\in\mathbb{R}}$ by first computing the inner products between $x(t)$ and the basis signals (Fourier transform) and then summing up the inner products times the basis (inverse Fourier transform). By analyzing the coordinates $X(\omega)$ of $x(t)$ with respect to this basis, one obtains a quantitative notion of "how much" of each

This work was supported in part by the U.S. Army Research Laboratory and the U.S. Army Research Office under grant W911NF-04-1-0221. Author E-mail: joebob@ou.edu, chuong@ou.edu, yeary@ieee.org.

Targets and Backgrounds XII: Characterization and Representation,
edited by Wendell R. Watkins, Dieter Clement, Proc. of SPIE Vol. 6239,
62390D, (2006) · 0277-786X/06/$15 · doi: 10.1117/12.666340

basis signal is present in $x(t)$. This notion is valuable in a wide variety of applications, especially those that are concerned with linear translation invariant filtering.

However, the Fourier representation can fail to provide the simplest, most intuitive interpretation of the signal (with respect to human visual or auditory perception). This occurs in particular when the signal is nonstationary in the sense that the frequency content is changing over time, so that finite length Fourier transforms computed over different time intervals are significantly different from one another. A classic example of this is the chirp signal $x(t) = \cos(t^2)$. If this signal is played through a loudspeaker, the human listener perceives a single tone with a fixed amplitude and a frequency or pitch that increases linearly over time. In the Fourier theory this signal must be represented using the only building blocks that are available, which are sinusoids of fixed amplitude and fixed frequency. Thus, in the Fourier representation, the nonstationary structure of the chirp signal must be represented by constructive and destructive interference between a large number of *stationary* sinusoids.

One motivation for performing time-frequency analysis is to obtain a representation of the chirp signal that better agrees with human hearing, so that it can be analyzed intuitively as a single sinusoid with a fixed amplitude and a linearly increasing frequency. The short-time Fourier transform (STFT) takes a step in this direction by providing a sequence of Fourier transforms $X(\omega, t)$ computed over finite length windows referenced to the time parameter t. Ideally, it would possible to consider the limit of the STFT as the window vanishes in order to obtain well posed notions of the instantaneous spectrum and amplitude. Then, the chirp could be represented unambiguously as a single tone (or component) $x(t) = a(t)\cos[\varphi(t)]$ with constant instantaneous amplitude $a(t) = 1$, quadratic instantaneous phase $\varphi(t) = t^2$, and linear instantaneous frequency $\varphi'(t) = 2t$.

AM-FM image models[1–3] arise from the extension of these notions into multiple dimensions. Suppose we are given a locally narrowband multidimensional signal $s : \mathbb{R}^n \to \mathbb{R}$. The objective is to find a positive semidefinite instantaneous amplitude or AM function $a(\mathbf{x})$ and an instantaneous frequency or FM function $\nabla\varphi(\mathbf{x})$ so that the signal $s(\mathbf{x})$ can be written as

$$s(\mathbf{x}) = a(\mathbf{x})\cos[\varphi(\mathbf{x})]. \tag{1}$$

An immediate problem is that the AM and FM function fail to admit well-posed definitions because, for any given $s(\mathbf{x})$, there exist uncountably infinitely many pairs of functions $a(\mathbf{x})$ and $\varphi(\mathbf{x})$ that provide equality in Eq. (1). Nevertheless, in many cases our intuition will deliver powerful notions of what the instantaneous amplitude and frequency *should be*, just as was the case with the chirp signal considered above. The multidimensional Teager-Kaiser operator and associated energy separation algorithm described in Ref. 2 provide one approach for associating an intuitively satisfying pair of AM and FM functions with the real-valued signal $s(\mathbf{x})$.

An alternative approach that we will adopt throughout this paper is to add an imaginary part $jq(\mathbf{x})$ to the real signal $s(\mathbf{x})$ to obtain a complex-valued signal $z(\mathbf{x}) = s(\mathbf{x}) + jq(\mathbf{x})$. The signal $z(\mathbf{x})$ may then be modeled according to $z(\mathbf{x}) = a(\mathbf{x})\exp[j\varphi(\mathbf{x})]$ where the AM function $a(\mathbf{x})$ and FM function $\nabla\varphi(\mathbf{x})$ are *unique*. These modulating functions may be computed directly from the values of the complex signal $z(\mathbf{x})$ using the exact demodulation algorithms[1]

$$\nabla\varphi(\mathbf{x}) = \mathrm{Re}\left[\frac{\nabla z(\mathbf{x})}{jz(\mathbf{x})}\right], \tag{2}$$

$$a(\mathbf{x}) = |z(\mathbf{x})|. \tag{3}$$

Thus we see that adding an imaginary part $jq(\mathbf{x})$ to the real signal $s(\mathbf{x})$ is precisely equivalent to selecting a particular pair of AM and FM functions to associate with $s(\mathbf{x})$. Therefore, the specific method of choosing $q(\mathbf{x})$ is of considerable importance. Let $\mathbf{x} = [x_1\ x_2\ \ldots\ x_n]^T$ and let $\mathbf{e}_1 = [1\ 0\ 0\ \ldots\ 0]^T$. Throughout this paper, we will take $q(\mathbf{x}) = \mathcal{H}[s(\mathbf{x})]$, where $\mathcal{H}$ is the directional (or *partial*) multidimensional Hilbert transform[1,3–5]

$$\mathcal{H}[s(\mathbf{x})] = \text{p.v.}\ \frac{1}{\pi}\int_{\mathbb{R}} s(\mathbf{x} - \xi\mathbf{e}_1)\frac{d\xi}{\xi} \tag{4}$$

and where the integral in Eq. (4) is interpreted as a Cauchy principle value. As described in Refs. 3 and 4, this provides a definition of the complex signal

$$z(\mathbf{x}) = s(\mathbf{x}) + j\mathcal{H}[s(\mathbf{x})] \tag{5}$$

that shares many of the attractive properties of the 1D analytic signal.[6–9]

2.1. Multi-Component Models

The complex model in Eq. (5) and demodulation algorithms in Eqs. (2) and (3) tend to deliver an intuitively appealing pair of AM and FM functions provided that the signal $s(\mathbf{x})$ is locally coherent in the sense that it is approximately sinusoidal over *sufficiently small* neighborhoods. However, for signals that are not locally coherent, such as, e.g., a sum of two or more pure sinusoids or of two or more chirps, the approach based on Eqs. (2), (3), and (5) will generally fail to provide a meaningful interpretation of the modulated signal structure.[8,10] In actuality, the types of complicated signals encountered in real-world engineering applications such as, e.g., typical infrared scenes, are *never* expected to be locally coherent. It therefore becomes necessary to model the signal of interest $s(\mathbf{x})$ not as a single AM-FM function as in Eq. (1), but rather as a sum of multiple *locally coherent* AM-FM functions given by

$$s(\mathbf{x}) = \sum_{k=1}^{K} a_k(\mathbf{x}) \cos[\varphi_k(\mathbf{x})] = \sum_{k=1}^{K} s_k(\mathbf{x}). \tag{6}$$

By linearity of the Hilbert transform in Eq. (4), it follows immediately that

$$z(\mathbf{x}) = s(\mathbf{x}) + j\mathcal{H}[s(\mathbf{x})] = \sum_{k=1}^{K} a_k(\mathbf{x}) \exp[j\varphi_k(\mathbf{x})] = \sum_{k=1}^{K} z_k(\mathbf{x}). \tag{7}$$

Because the demodulation algorithms given in Eq. (2) and (3) are nonlinear, the components $z_k(\mathbf{x})$ in Eq. (7) must be isolated from one another on a spatially local basis prior to demodulation. This is typically accomplished by applying a multiband bank of linear translation invariant filters.[1–3] In order to avoid demodulation errors due to cross component interference, it is necessary for the filters to be spectrally localized. However, they must also be spatially localized to avoid averaging out the important spatial texture structure of the signal. In view of these conflicting design goals, the Gabor filters, which in the continuous case uniquely optimize the uncertainty principle lower bound on joint spatio-spectral localization,[6,11] have frequently been employed. The impulse response of a unit L^2-norm isotropic unity-variance Gabor filter at baseband is given by

$$g(\mathbf{x}) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{4}\mathbf{x}^T\mathbf{x}\right). \tag{8}$$

The frequency response of this filter is an n-dimensional isotropic Gaussian centered about the frequency origin. By adding frequency scaling and frequency translation, we obtain the bandpass impulse response

$$g_k(\mathbf{x}) = \frac{1}{\sigma_k\sqrt{2\pi}} \exp\left(-\frac{1}{4\sigma_k^2}\mathbf{x}^T\mathbf{x}\right) \exp\left(j\boldsymbol{\Omega}_k^T\mathbf{x}\right) \tag{9}$$

with center frequency $\boldsymbol{\Omega}_k \in \mathbb{R}^n$ and half-peak radial octave bandwidth given by

$$B = \log_2\left(\frac{|\boldsymbol{\Omega}_k| + \frac{\sqrt{\ln 2}}{\sigma_k}}{|\boldsymbol{\Omega}_k| - \frac{\sqrt{\ln 2}}{\sigma_k}}\right). \tag{10}$$

The frequency response of the bandpass filter is given by the isotropic Gaussian

$$G_k(\boldsymbol{\Omega}) = \mathcal{F}[g_k(\mathbf{x})] = 2\sqrt{2\pi}\sigma_k \exp\left[-\sigma_k^2(\boldsymbol{\Omega} - \boldsymbol{\Omega}_k)^T(\boldsymbol{\Omega} - \boldsymbol{\Omega}_k)\right]. \tag{11}$$

By construction, the spectrum of the complex signal $z(\mathbf{x})$ in Eq. (7) is supported only in 2^{n-1} frequency orthants where it is given by $Z(\boldsymbol{\Omega}) = 2S(\boldsymbol{\Omega})$. Consequently, in designing the multiband filterbank it is sufficient to tile half the frequency space with a tesselation of filters $G_k(\boldsymbol{\Omega})$. The frequency response of a typical 33-channel 2D filterbank is shown in Fig. 1(b), where the filters cover quadrants I and IV of the 2D frequency plane. The filters are arranged in a polar tesselation comprising four filters at each of eight orientations. Each filter has a unity radial octave bandwidth and any group of four adjacent filters intersect in a single frequency where all four

are at half of peak response. The structure of the filterbank *defines* the decomposition of the signal $z(\mathbf{x})$ into components $z_k(\mathbf{x})$ as indicated in Eq. (7). For each filterbank channel $G_k(\mathbf{\Omega})$, we take

$$y_k(\mathbf{x}) = z(\mathbf{x}) * g_k(\mathbf{x}) \approx z_k(\mathbf{x}) * g_k(\mathbf{x}), \tag{12}$$

where the symbol "$*$" indicates linear convolution and, by construction, the channel response $y_k(\mathbf{x})$ is dominated by signal component $z_k(\mathbf{x})$. While the approximation error inherent in Eq. (12) prohibits the possibility of perfectly reconstructing the signal from the computed modulations obtained from the filterbank channel responses, this is usually not a concern in analysis-only applications such as those considered here.

The modulating functions $a_k(\mathbf{x})$ and $\nabla\varphi_k(\mathbf{x})$ of the individual components $z_k(\mathbf{x})$ in Eq. (7), are estimated by applying a novel family of quasi-eigenfunction approximations (QEA's)[1] to factor the effects of the filter $G_k(\mathbf{\Omega})$ out of the response $y_k(\mathbf{x})$. This provides a closed-form expression for the response in terms of the input modulations that can be solved to arrive at the *filtered* demodulation algorithms

$$\nabla\varphi_k(\mathbf{x}) \approx \mathrm{Re}\left[\frac{\nabla y_k(\mathbf{x})}{j y_k(\mathbf{x})}\right], \tag{13}$$

$$a_k(\mathbf{x}) \approx \left|\frac{y_k(\mathbf{x})}{G_k[\nabla\widehat{\varphi}_k(\mathbf{x})]}\right|, \tag{14}$$

which may be applied to each channel response on a pointwise basis.

2.2. Discrete Implementation

To apply the approach given in Section 2.1 to 2D digital images, we obtain the discrete filterbank by frequency sampling of the continuous design. For an $N \times N$ discrete image $s : \mathbb{Z}^2 \to \mathbb{R}$, let $S(u,v) = \mathrm{DFT}[s(\mathbf{x})]$. The DFT domain discretization of the Hilbert transform in Eq. (4) is given by[3,12]

$$\mathcal{H}(u,v) = \begin{cases} -j, & u = 1,2,\ldots,\frac{N}{2}-1, \\ j, & u = \frac{N}{2}+1, \frac{N}{2}+2,\ldots,N-1, \\ -j, & u = 0,\ v = 1,2,\ldots,\frac{N}{2}-1, \\ -j, & u = \frac{N}{2},\ v = 1,2,\ldots,\frac{N}{2}-1, \\ j, & u = 0,\ v = \frac{N}{2}+1, \frac{N}{2}+2,\ldots,N-1, \\ j, & u = \frac{N}{2},\ v = \frac{N}{2}+1, \frac{N}{2}+2,\ldots,N-1, \\ 0, & \text{otherwise.} \end{cases} \tag{15}$$

The complex image $z(m,p)$ is then obtained by $z(m,p) = \mathrm{DFT}^{-1}[S(u,v)\mathcal{H}(u,v)]$. For the discrete FM function $\nabla\varphi_k(m,p)$, which contains the samples of $\nabla\varphi_k(\mathbf{x})$, let $\nabla\varphi_k(m,p) = [u_k(m,p)\ v_k(m,p)]^T$. The QEA[1] may then be applied to obtain the following discretization of Eq. (13):

$$|u_k(m,p)| \approx \arccos\left[\frac{y_k(m+1,p)+y_k(m-1,p)}{2y_k(m,p)}\right], \tag{16}$$

$$\mathrm{sgn}\, u_k(m,p) \approx \mathrm{sgn}\arcsin\left[\frac{y_k(m+1,p)-y_k(m-1,p)}{2jy_k(m,n)}\right], \tag{17}$$

$$|v_k(m,p)| \approx \arccos\left[\frac{y_k(m,p+1)+y_k(m,p-1)}{2y_k(m,p)}\right], \tag{18}$$

$$\mathrm{sgn}\, v_k(m,p) \approx \mathrm{sgn}\arcsin\left[\frac{y_k(m,p+1)-y_k(m,p-1)}{2jy_k(m,p)}\right]. \tag{19}$$

The filtered amplitude algorithm of Eq. (14) may then be sampled according to

$$a_k(m,p) \approx \frac{|y_k(m,p)|}{|G_k[u_k(m,p), v_k(m,p)]|}. \tag{20}$$

2.3. Dominant Modulations

For each pixel in a discrete image $s(m,p)$, it is of interest to obtain a modulation domain feature vector by extracting the AM and FM estimates $a_k(m,p)$ and $\nabla\varphi_k(m,p)$ given in Eqs. (16)-(19) and (20) corresponding to the component $z_k(m,p)$ that dominates the instantaneous image spectrum in a local neighborhood about the pixel. To this end, we compute at each pixel a channel selection criterion given by

$$\Gamma_k(m,p) = \frac{|y_k(m,p)|}{\max_{u,v}|G_k(u,v)|}, \tag{21}$$

where $1 \leq k \leq K$. The dominant modulations are then given by $\{a_D(m,p), \nabla\varphi_D(m,p)\}$, where $D = \text{argmax}_k\, \Gamma_k(m,p)$. With the QEA, it may be shown that

$$\Gamma_k(m,p) \approx |a_k(m,p)| \frac{|G_k[\nabla\varphi_k(m,p)]|}{\max_{u,v}|G_k(u,v)|}. \tag{22}$$

Thus, choosing the dominant component as the one that dominates the response of the channel that maximizes Γ_k in Eq. (21) tends to select the component $z_k(m,p)$ that has the largest amplitude, and, moreover, to extract estimates of the AM and FM functions of this component from the channel for which the estimated frequency vector lies nearest to the channel center frequency. This approach ensures that, even when the input in fact contains only a few components so that most channels are driven primarily by noise and cross component interference, the approach based on Eq. (21) will tend to select the strongest component and extract the dominant modulations from the channel for which the signal-to-noise ratio (SNR) is best.

3. APPLICATION TO INFRARED TARGETS AND BACKGROUNDS

Visually, the dominant modulations characterize and quantify the dominant local texture structure of an image. The dominant amplitude $a_D(m,p)$ captures the local texture contrast. The local texture orientation is given by $\theta(m,p) = \arg\nabla\varphi_D(m,p)$, while the spatial granularity is quantified by $R(m,p) = |\nabla\varphi_D(m,p)|$. We consider two main questions in this paper. First, does typical infrared imagery possess sufficient texture structure to admit a meaningful characterization in the modulation domain? Second, is the modulation domain feature vector potentially useful for enhancing infrared target-background class separability? In the modulation domain feature vector, we generally expect naturally occurring backgrounds to exhibit a lower degree of organization or coherency than, e.g., targets such as military vehicles. However, we expect both man made and natural clutter to potentially be highly organized.

Because the richly textured surface patterns characteristic of visible images tend to occur to a much lower degree in infrared imagery, the measurement of dominant modulations in an infrared image is challenging and highly sensitive to the filterbank structure. Whereas the generic filterbank of Fig. 1(b) may be used to compute modulations for a wide variety of visible images for applications such as image segmentation, we have found that it is less likely to be successful when applied to typical infrared images, as illustrated in Fig. 1. A typical scene is shown in Fig. 1(a).[13] The generic filterbank was applied to compute the dominant AM function given in Fig. 1(c). A reconstruction of the dominant AM-FM image component appears in Fig. 1(d). Each patch or region in Fig. 1(d) shows the nonstationary quasi-sinusoid that best agrees with the dominant modulations. The patches themselves arise as a consequence of the fact that different image components $z_k(m,p)$ are dominant in different image regions. Here, we see that the dominant modulations have been determined almost entirely by the highly structured background with the result that the modulation domain features utterly fail to reveal the target. By contrast, the much less dense filterbank structure shown in Fig. 1(e) effectively captures the heated tracks of the vehicle, which appear prominently in the computed dominant AM function of Fig. 1(f) and dominant component reconstruction of Fig. 1(g).

To effectively measure the dominant modulations of infrared images, we consider that the modulation domain feature vector will be used as an auxiliary observation to improve the performance of a standard target tracker such as an extended Kalman filter or particle filter in an automatic or assisted target recognition system. Thus, we assume that there is an independent detection process, that a target designation is available in the initial

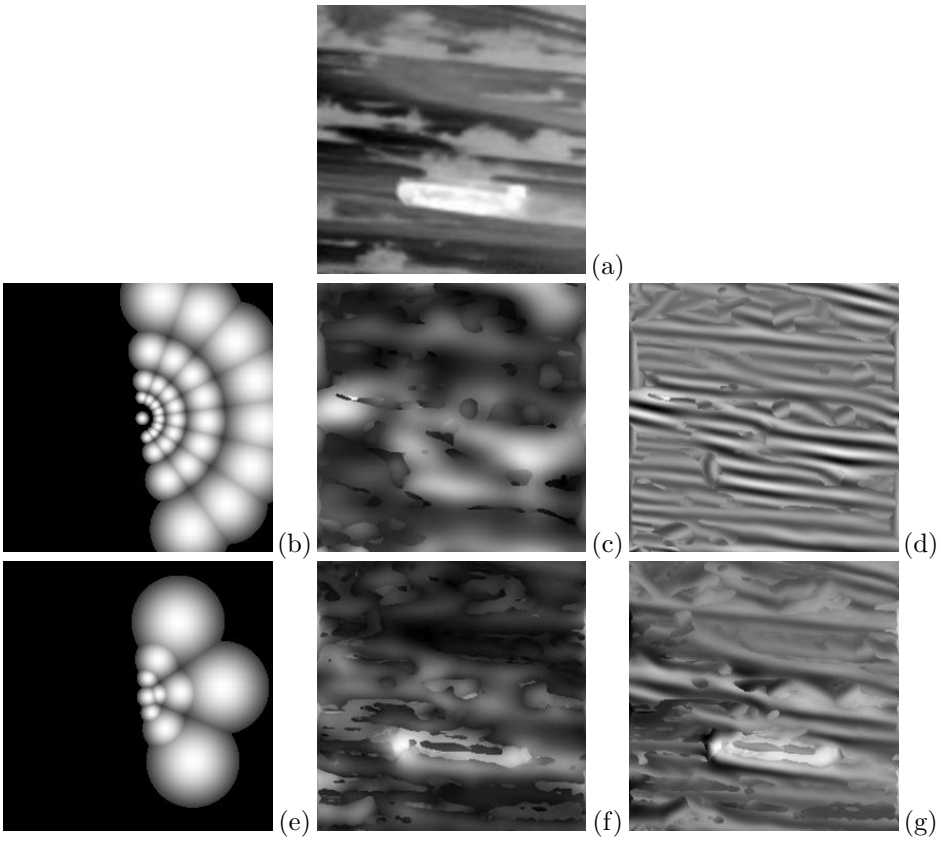


Figure 1. Example illustrating the sensitivity of the computed modulations to the filterbank structure when computing infrared AM-FM models. (a) Typical scene.[13] (b) 33-channel filterbank of the type that is usually successful for computing AM-FM models of visible images. (c) Dominant AM function obtained with the filterbank of (b). (d) Reconstruction of the dominant AM-FM image function obtained with the filterbank of (b). (e) The filters in this bank have nearly twice the typical bandwidth and a much less dense spacing than the typical bank. (f),(g) Dominant AM function and reconstructed dominant image component obtained with the filterbank of (e).

frame of any image sequence, and that a predicted track gate is available in subsequent frames. Initially, we make use of a filterbank with the typical structure as shown in Fig. 1(b).

For practical sequences depicting an extended target, the designated target signature in the first frame generally spans multiple pixels where several different filterbank channels optimize the selection criterion of Eq. 21. We initially choose a channel that lies centrally within the locus of these optimizing channels. We construct a Gabor filter $\mathcal{G}_0(u,v)$ having a center frequency identical to that of the chosen filter, but with a fixed linear bandwidth of 0.1 cycles per image. We demodulate the response of this filter to extract the modulation domain feature vector of the target over the pixels that lie in the initial track gate. For each subsequent frame, the target filter $\mathcal{G}_{k+1}(u,v)$ is adaptively generated from $\mathcal{G}_k(u,v)$ by updating the radial and angular components of the center frequency as follows. Let r_k and ϑ_k be the radial frequency and orientation of center frequency of $\mathcal{G}_k(u,v)$. Similarly, let $\overline{R}_k$ be the mean radial frequency $|\nabla\varphi(m,p)|$ and $\overline{\theta}_k$ be the mean orientation $\arg\nabla\varphi(m,p)$ of the target signature measured in frame k. The center frequency of $\mathcal{G}_{k+1}(u,v)$ is given by

$$r_{k+1} = r_k - (r_k - \overline{R}_k)\Delta_r \tag{23}$$

$$\vartheta_{k+1} = \vartheta_k - (\vartheta_k - \overline{\theta}_k)\Delta_\vartheta, \tag{24}$$

where Δ_r and Δ_ϑ are adaptation step sizes that limit misadjustment due to noise and cross component interference. For sequences depicting rapidly evolving kinematics, the step sizes must be designed to balance noise rejection against the adaptation speed of the filter.

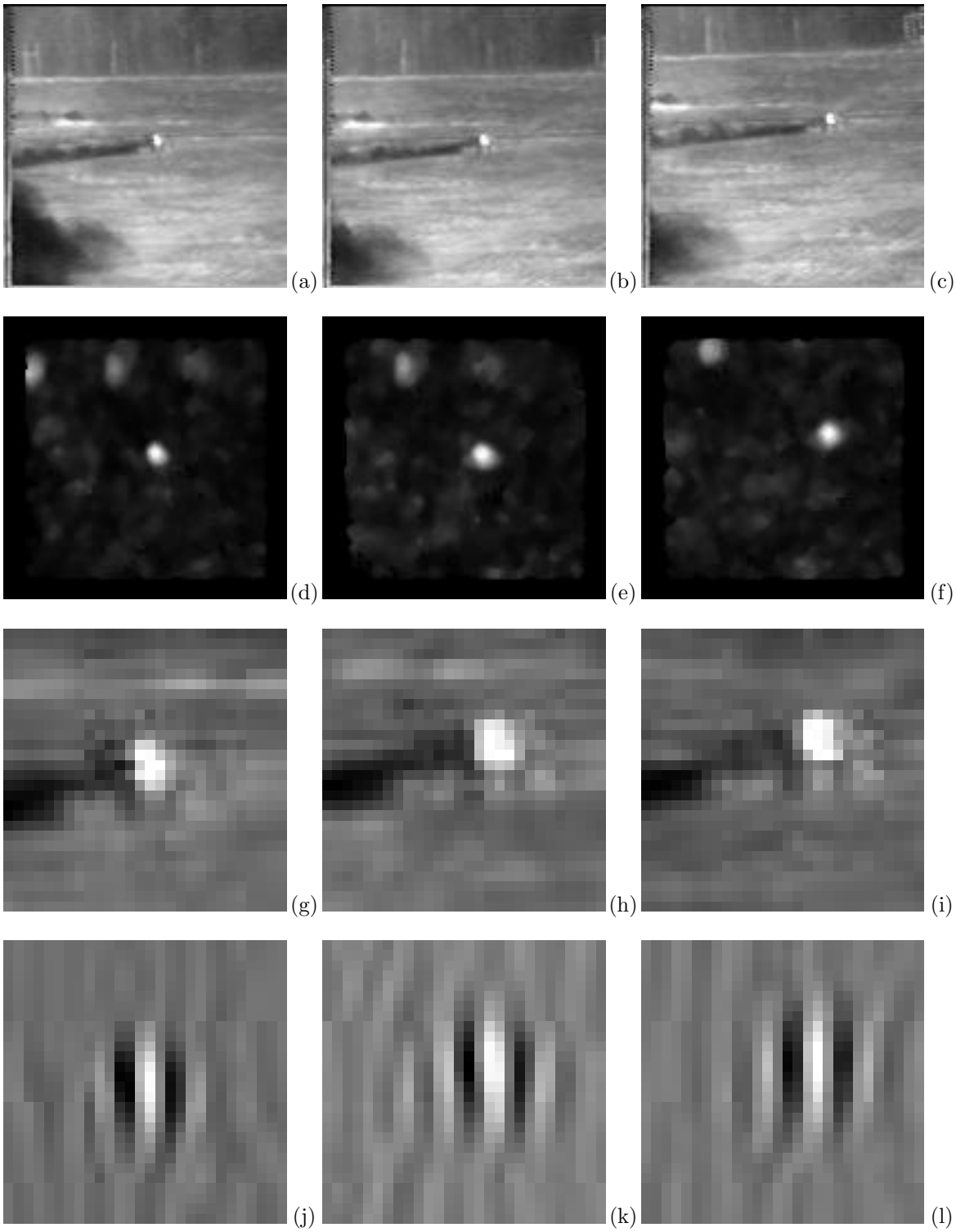


Figure 2. Example showing small scale adaptation of the target filter with excellent clutter suppression in the computed dominant AM function. (a)-(c) Original frames. (d)-(f) Dominant AM function computed from the target filter. (g)-(i) Detail views of the target signature from the original frames. (j)-(l) Detail views of the reconstructed dominant AM-FM component computed from the target filter.

4. EXAMPLES

In this section, we present two examples where the adaptive Gabor filter described in Section 3 is used to extract modulation domain target features from AMCOM longwave (8 – 12 μm) infrared missile approach sequences. The kinematics in these sequences tend to be relatively mild, and we typically take $0.01 \leq \Delta_r \leq 0.02$ and $0.1 \leq \Delta_\vartheta \leq 0.4$ in Eq. (23) and (24). Three frames from the first sequence are shown in Fig. 2(a)-(c). These are frames 1, 5, and 9 of the sequence, which demonstrate small scale adjustments of the filter $\mathcal{G}_k$ over short time intervals (the entire sequence contained 30 frames). The target is a truck which is in motion. The dominant AM function computed over each entire image using the target filter is given in Fig. 2(d)-(f), where we see substantial suppression of the background structure and amplification of the target. Detail views of the target extracted from the unprocessed frames are shown in Fig. 2(g)-(i), while corresponding reconstructions of the dominant AM-FM component extracted by the adaptive Gabor filter are given in Fig. 2(j)-(l). Although we do not have detailed ground truth information for this sequence, it appears that wheels are visible below the brightest part of the target signature, and the horizontal granularity of these structures dominates the modulation domain interpretation of the target signature. The distant structure at the top of the original frames appears to be a tree line and it is interesting to note that in places the spacing of the trees matches that of the wheels, leading to a nontrivial response in the AM images of Fig. 2(d)-(f). In this example, the filter orientation ϑ_k adapted by 0.5° and the radial center frequency 0.4 cycles per image. The adaptation step sizes were set to $\Delta_r = 0.01$ and $\Delta_\vartheta = 0.1$.

As shown in Fig. 3, the second example shows dramatic adaptation of the filter $\mathcal{G}_k$ over a much longer sequence. Fig. 3(a)-(d) show, respectively, frames 50, 80, 90, and 99 out of a total of 100 frames. In this case, the magnification of the target changes substantially through the sequence. The adaptation step sizes were set to $\Delta_r = 0.02$ and $\Delta_\vartheta = 0.4$. The filter radial center frequency decreased monotonically from 12.8 to 5.5 cycles per image, while the orientation varied between -2.8° and 70°. The dominant AM functions extracted from the target filter are given in Fig. 3(e)-(h) and once again show substantial suppression of the background structure. Reconstructions of the dominant AM-FM image component extracted from the target filter are given in Fig. 3, where the steady decrease in the dominant radial frequency over the course of the sequence can be seen. In order to provide a qualitative assessment of the usefulness of the computed AM-FM models for enhancing target-background separability, we constructed scatter plots of the modulation domain feature vectors for frames 50 and 90 of this sequence. The scatter plots are given in Fig. 4(a) and (b), respectively, where the target pixels are shown as red circles and the background pixels are shown as blue crosses. The AM and FM functions for the target were extracted from the adaptive Gabor filter $\mathcal{G}_k$ over the pixels contained in the track gate. For the remaining (background) pixels, the dominant modulations were extracted from the filterbank of Fig. 1(a) using the channel selection criterion given by Eq. (21). These plots demonstrate the ability of the modulation domain representation to pull the target up and out of the background, despite the fact that the raw pixel values from the two classes are substantially overlapped. It is also clear from Fig. 4(a) and (b) that the amplitude axis of the feature space provides by far the most powerful class separability information in this sequence.

5. CONCLUSION

We applied the theory of AM-FM image modeling to construct modulation domain feature vectors for infrared targets and backgrounds. For each pixel, the feature vector comprises a scalar amplitude modulation value $a(m,p)$ and a two-element instantaneous frequency vector $\nabla\varphi(m,p)$. These vectors can be used as auxiliary inputs to enhance the performance of track processors and ATR systems that would otherwise operate on the pixel values alone. Under the assumption that predicted track gates and an initial target designation are available as side information, we constructed adaptive Gabor filters for measuring the target modulating functions. The dominant background modulations were computed using a fixed filterbank structure with a selection criterion for determining the channel that contains the dominant background AM-FM component. Our results show that typical infrared imagery does possess sufficient local texture structure to admit a useful and meaningful characterization in the modulation domain. In addition, the feature vector scatter plots of Fig. 4 demonstrate qualitatively that computed AM-FM models have the potential to offer powerful class separability information for discriminating infrared targets and backgrounds.

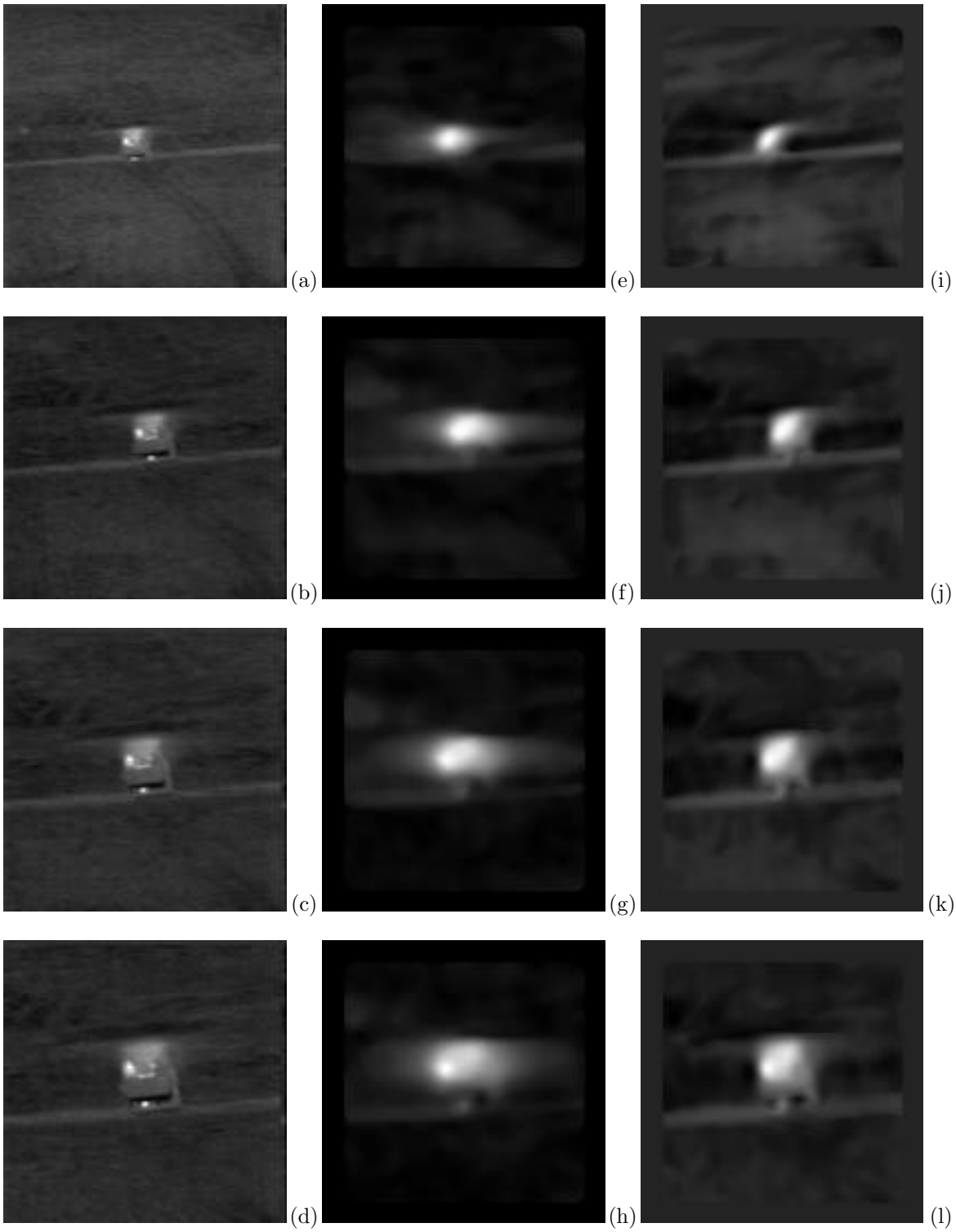
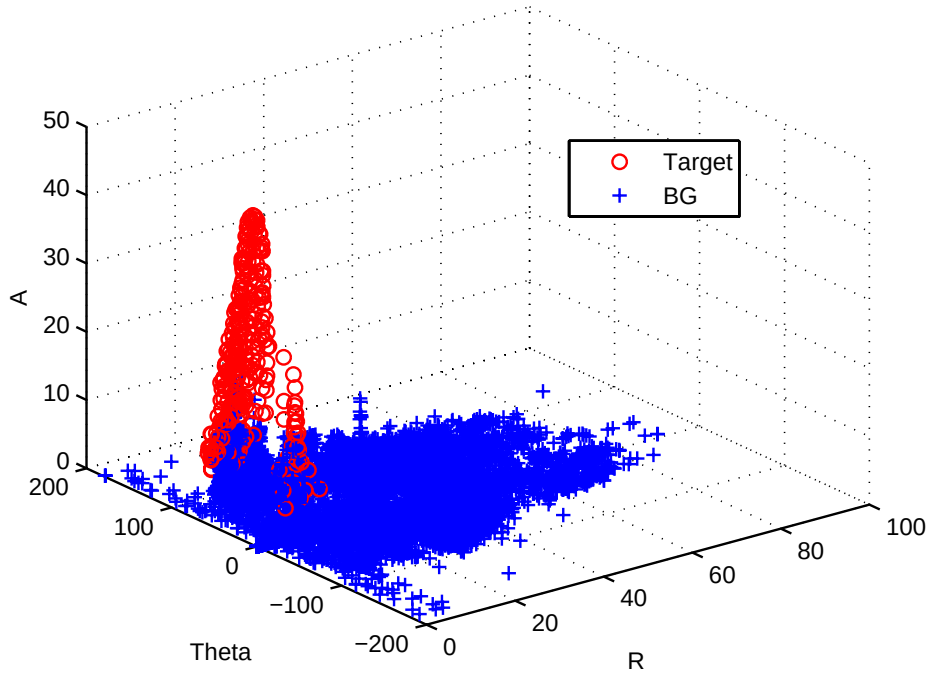
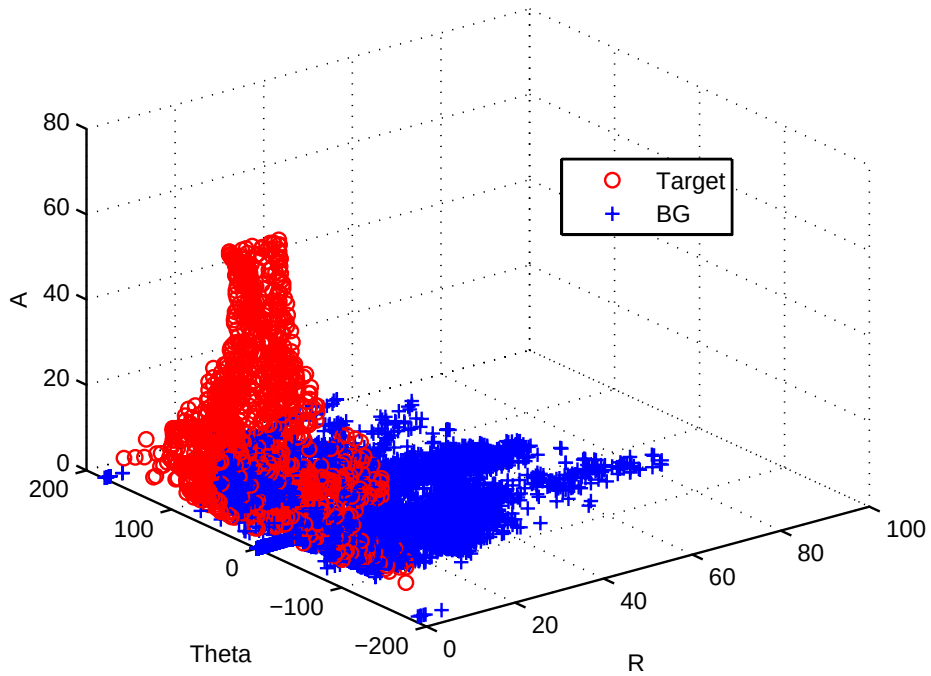


Figure 3. Example showing large scale adaptation over a longer sequence. (a)-(d) Original frames. (e)-(h) Dominant AM functions computed from the target filter. (i)-(l) Reconstructions of the dominant AM-FM components extracted from the target filter.



(a)



(b)

Figure 4. Scatter plots of the modulation domain feature vectors for frames 50 and 90 of the sequence depicted in Fig. 3. (a) Scatter plot for frame 50. (b) Scatter plot for frame 90. Target pixels are shown as red circles, while background pixels are shown as blue crosses.

ACKNOWLEDGMENTS

The AMCOM missile closure sequences in Fig. 2 and 3 were provided through the Johns Hopkins University Center for Imaging Science under ARO DAAH049510494. This work was supported in part by the U.S. Army Research Laboratory and the U.S. Army Research Office under grant W911NF-04-1-0221.

REFERENCES

1. J. P. Havlicek, D. S. Harding, and A. C. Bovik, "Multidimensional quasi-eigenfunction approximations and multicomponent AM-FM models," *IEEE Trans. Image Proc.* **9**, pp. 227–242, Feb. 2000.
2. P. Maragos and A. C. Bovik, "Image demodulation using multidimensional energy separation," *J. Opt. Soc. Amer. A* **12**, pp. 1867–1876, Sep. 1995.
3. J. P. Havlicek, P. C. Tay, and A. C. Bovik, "AM-FM image models: Fundamental techniques and emerging trends," in *Handbook of Image and Video Processing*, A. C. Bovik, ed., pp. 377–395, Elsevier Academic Press, Burlington, MA, 2nd ed., 2005.
4. J. P. Havlicek, J. W. Havlicek, and A. C. Bovik, "The analytic image," in *Proc. IEEE Int'l. Conf. Image Proc.*, (Santa Barbara, CA), Oct. 26-29, 1997.
5. J. P. Havlicek, J. W. Havlicek, N. D. Mamuya, and A. C. Bovik, "Skewed 2D Hilbert transforms and computed AM-FM models," in *Proc. IEEE Int'l. Conf. Image Proc.*, pp. 602–606, (Chicago, IL), Oct. 4-7, 1998.
6. D. Gabor, "Theory of communication," *J. Inst. Elect. Eng. London* **93**(III), pp. 429–457, 1946.
7. J. Ville, "Théorie et applications de la notation de signal analytique," *Cables et Transmission* **2A**, pp. 61–74, 1948. translated from the French in I. Selin, "Theory and applications of the notion of complex signal," Tech. Rept. T-92, The RAND Corporation, Santa Monica, CA, Aug. 1958.
8. L. Cohen, *Time-Frequency Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1995.
9. A. Papoulis, *The Fourier Integral and its Applications*, McGraw-Hill, New York, 1962.
10. D. Wei and A. C. Bovik, "On the instantaneous frequencies of multicomponent AM-FM signals," *IEEE Signal Proc. Let.* **5**, pp. 84–86, Apr. 1998.
11. 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* **2**, pp. 1160–1169, Jul. 1985.
12. V. Čížek, "Discrete Hilbert transform," *IEEE Trans. Audio, Electroacoust.* **18**, pp. 340–343, Dec. 1970.
13. http://www.unt.edu/resource/02invisiblefeature.htm, visited March, 2006.