MODULATION DOMAIN FEATURES FOR DISCRIMINATING INFRARED TARGETS AND BACKGROUNDS

Chuong T. Nguyen and Joseph P. Havlicek

School of Electrical and Computer Engineering
University of Oklahoma, Norman, OK 73019 USA

ABSTRACT

For the first time, we compute modulation domain features for infrared targets and backgrounds, including dominant modulations that characterize the local texture contrast, orientation, and granularity. We present a practical computational approach and introduce a new FM algorithm designed to reduce the approximation errors characteristic of many existing discrete techniques. By performing experiments against actual FLIR approach sequences, we verify that typical IR imagery does indeed possess sufficient texture structure for effective modulation domain characterization. We demonstrate qualitatively that the modulation domain features can significantly enhance target-background class separability relative to pixel domain features.

Index Terms— Amplitude modulation, frequency modulation, infrared imaging, infrared tracking, object recognition

1. INTRODUCTION

The problem of detecting military targets in forward-looking infrared (FLIR) imagery has been studied extensively. As is well-known, it is an extremely challenging problem (see, e.g., [1] and the references therein) due to the confluence of several factors, including both the complex nature of the thermal scenes typical of the modern battlespace and the relative immaturity of the materials used in fabricating infrared sensors compared to their visible wavelength counterparts. In particular, FLIR images frequently exhibit weak signal-to-noise ratio and strong clutter, where the term *clutter* refers generically to any structures in the image that arise from sources other than the target. Often, the radiometric signature of the background can be as strong as or stronger than the actual target signature, and it is not uncommon for the clutter variance to also be stronger than that of the target. The inherent difficulty of this problem is borne out by the vast body of literature that has been devoted to it.

In this paper, we introduce a radically new approach. For the first time, we consider computing low-level target and background features in the *modulation domain* as opposed to traditional pixel domain and Fourier domain features. Our central hypothesis, which is supported by the preliminary experimental results given in Section 4, is that, in the modulation domain, military targets should exhibit a significantly higher degree of organization or *local coherency* [2, 3] compared to naturally occurring clutter backgrounds. Our results strongly suggest that this characteristic can be used to advantage to improve target detection performance in a variety of FLIR signal processors by augmenting their feature vectors with modulation domain data.

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.

2. MODULATION DOMAIN IMAGE MODEL

Multidimensional AM-FM models seek to represent an image $t : \mathbb{R}^n \rightarrow \mathbb{R}$ as a sum of quasi-sinusoidal components of the form $a(\mathbf{x})\cos[\varphi(\mathbf{x})]$, where the AM function (i.e., instantaneous amplitude) $a(\mathbf{x})$ is positive semidefinite and the FM function (i.e., instantaneous frequency) $\nabla\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$. This is motivated by a compelling body of psychophysical evidence indicating that, quantitatively, these modulating functions are intimately related to visual perception in humans and other mammals.

For a real-valued image, the modulating functions are ambiguous in the sense that they cannot be uniquely defined. Although excellent techniques such as the Teager-Kaiser operator [4] exist for estimating the modulating functions directly from the real values of an image, we prefer to disambiguate the problem by adding to $t(\mathbf{x})$ an imaginary component equal to the multidimensional Hilbert transform given in [5]. We model the resulting complex-valued image $z(\mathbf{x})$ according to

$$z(\mathbf{x}) = t(\mathbf{x}) + j\mathcal{H}[t(\mathbf{x})] = \sum_{k=1}^{K} a_k(\mathbf{x})\exp[j\varphi_k(\mathbf{x})], \quad (1)$$

where $t(\mathbf{x})$ is interpreted as a sum of K AM-FM components. This model has been used successfully to treat textured images in particular, where a_k admits an interpretation as the local texture contrast of the k'th component, whereas $\nabla\varphi_k$ characterizes the local texture orientation and granularity. Discrete algorithms for estimating the modulating functions from the samples of $z(\mathbf{x})$ were given in [3, 4, 6].

The fundamental question we address in this paper is the following: do infrared targets and backgrounds possess sufficient texture structure that (1) can be applied to enhance target-background class separability in FLIR imagery? While a definitive answer to this question is beyond the scope of this paper, we will present compelling examples in Section 4 that strongly suggest, albeit in a qualitative sense, that the answer is *yes*. This finding is important because it indicates that modulation domain features have the potential to make a significant impact on the notoriously difficult infrared target detection problem. In Section 3, we describe techniques for computing the modulation domain features.

3. FEATURE SPACE COMPUTATION

Since all practical discrete AM-FM demodulation algorithms are nonlinear, it is generally necessary to apply some type of filtering to isolate the components in (1) from one another on a spatially local basis prior to estimating the individual modulating functions. The filters must also be well localized in frequency to resolve closely spaced components and avoid cross-component interference; consequently, the complex-valued Gabor filters, which admit optimal joint time-frequency localization, are a popular choice. Hence, we assume a bank of Gabor filters with impulse responses $g_k(\mathbf{x})$ and frequency responses $G_k(\mathbf{\Omega})$, where $k \in [1, K]$ and $\mathbf{x}, \mathbf{\Omega} \in \mathbb{R}^n$.

Let $z_k(\mathbf{x}) = a_k(\mathbf{x}) \exp[j\varphi_k(\mathbf{x})]$ and let $y_k(\mathbf{x})$ be the response of the kth filter, so that $y_k(\mathbf{x}) = z(\mathbf{x}) * g_k(\mathbf{x}) \approx z_k(\mathbf{x}) * g_k(\mathbf{x})$. Applying the quasi-eigenfunction approximations given in [2, 3], we obtain

$$y_k(\mathbf{x}) \approx z_k(\mathbf{x}) G[\nabla\varphi_k(\mathbf{x})], \tag{2}$$

which motivates the FM estimation algorithm [3]

$$\nabla\varphi_k(\mathbf{x}) \approx \mathrm{Re}\left[\frac{\nabla y_k(\mathbf{x})}{j y_k(\mathbf{x})}\right]. \tag{3}$$

The AM function $a_k(\mathbf{x})$ may then be obtained by

$$a_k(\mathbf{x}) \approx \left|\frac{y_k(\mathbf{x})}{G[\nabla\varphi_k(\mathbf{x})]}\right|. \tag{4}$$

In practice, the main problem with (3) and (4) is that additional approximation theory must be applied to obtain discrete versions of these algorithms, and this can result in nontrivial approximation errors if the modulating functions fail to be sufficiently smooth over one or more local spatial neighborhoods.[1] This problem is not unique to the techniques presented here and is also characteristic of, e.g., the discrete energy separation algorithms associated with the n-D Teager-Kaiser operator described in [4]. Consequently, we as well as others have recently investigated schemes for fitting continuous interpolants to the discrete image samples. With this approach it becomes possible to apply continuous algorithms such as (3) and (4) directly to the interpolating functions, where the results are subsequently restricted to the pixel lattice to obtain a discrete solution. For example, cubic tensor product splines were fit to the image phase in [6,7], while 1-D cubic splines were fit to the real-valued signal samples in [8]. Unfortunately, extending this idea to compute $\nabla\varphi_k(\mathbf{x})$ in (1) by direct continuous domain differentiation requires unwrapping the phase of $y_k(\mathbf{x})$ prior to performing the spline interpolation.

Here, we consider a new approach that entirely circumvents the difficult multidimensional phase unwrapping problem. Noting that $G[\nabla\varphi_k(\mathbf{x})]$ is both real and nonnegative, we let $\alpha_k(\mathbf{x}) = a_k(\mathbf{x}) G[\nabla\varphi_k(\mathbf{x})]$. From (2), we then have that $y_k(\mathbf{x}) \approx \alpha_k(\mathbf{x}) \exp[j\varphi_k(\mathbf{x})]$ and $\alpha_k(\mathbf{x}) \approx |y_k(\mathbf{x})|$, whereupon it follows that (3) may be expressed as

$$\nabla\varphi_k(\mathbf{x}) \approx \frac{1}{|y_k(\mathbf{x})|^2}\Big(\mathrm{Re}\,[y_k(\mathbf{x})]\,\mathrm{Im}\,[\nabla y_k(\mathbf{x})] - \mathrm{Im}\,[y_k(\mathbf{x})]\,\mathrm{Re}\,[\nabla y_k(\mathbf{x})]\Big). \tag{5}$$

We obtain estimates of the AM and FM functions of all K components in (1) at all pixels by first fitting the cubic tensor product splines described in [9, 10] to the real and imaginary parts of the complex images $y_k(\mathbf{x})$, then applying (5) followed by (4) to the K resulting spline models, and finally discretizing the obtained AM and FM functions to the pixel lattice.

At each each pixel $\mathbf{m} \in \mathbb{Z}^n$, we apply the dominant component analysis (DCA) computational paradigm given in [6] to extract dominant modulations $a_D(\mathbf{m})$ and $\nabla\varphi_D(\mathbf{m})$ corresponding to the AM-FM component that dominates the local structure of $t(\mathbf{x})$ at the pixel, where

$$D = \underset{k}{\mathrm{argmax}} \left|\frac{y_k(\mathbf{m})}{\max_{\mathbf{\Omega}} G_k(\mathbf{\Omega})}\right|. \tag{6}$$

The 3-D modulation domain feature space of the image $t(\mathbf{x})$ is then given by the locus of vectors $[A(\mathbf{m})\ R(\mathbf{m})\ \theta(\mathbf{m})]^T$, where $A(\mathbf{m}) = a_D(\mathbf{m})$, $R(\mathbf{m}) = |\nabla\varphi_D(\mathbf{m})|$, and $\theta(\mathbf{m}) = \arg\nabla\varphi_D(\mathbf{m})$.

4. EXAMPLE

We applied the techniques given in Section 3 to compute modulation domain feature vectors for a comprehensive set of AMCOM infrared missile closure sequences provided by the Johns Hopkins University Center for Imaging Science. This dataset consists of 49 longwave (LWIR) and midwave (MWIR) FLIR sequences, each containing about 100 frames. Each frame comprises 128×128 8-bit pixels.

Computation of the feature space is not terribly sensitive to the specific structure of the filterbank, provided that the frequency half-plane where the discrete complex image

[1] This notion of local coherency is is made precise in [3].

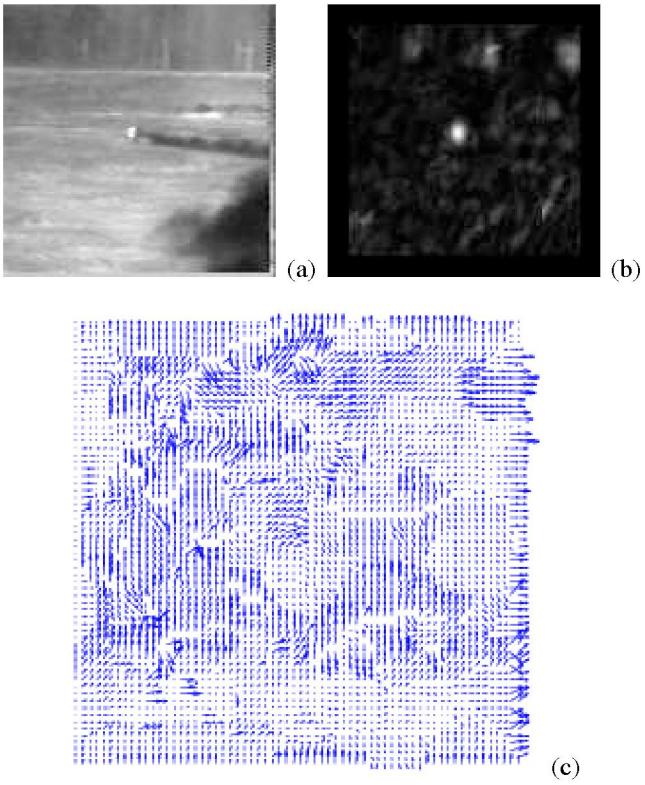

Fig. 1. Original image and dominant modulations. (a) Original FLIR image. (b) Dominant AM function $A(\mathbf{m})$. (c) Dominant FM function $\nabla\varphi_D(\mathbf{m})$ depicted as a needle diagram. Each arrow points in the direction $\theta(\mathbf{m})$ and has length proportional to $R(\mathbf{m})$.

$z(\mathbf{m})$ is supported is covered by a reasonably dense tesselation of jointly localized filters. For the AMCOM closure sequences we used a polar Gabor filterbank of the type described in [2, 11], where 32 constant-octave bandwidth channels were arranged along rays at eight orientations with four filters per ray. Each filter had a bandwidth of one octave, and the center frequencies along each ray followed a geometric progression with common ratio 1.8, with the lowest frequency filter being placed at a radial center frequency of 9.6 cycles per image (cpi). Adjacent filters intersected at half-peak, resulting in an angular ray spacing of 20.64°.

A typical frame from one of the LWIR sequences is shown in Fig. 1(a). Since our objective here is to demonstrate the utility of modulation domain features in discriminating infrared targets and backgrounds, and most certainly *not* to propose a complete target detection algorithm, we used *a priori* ground truth information about the target location and applied (6) to determine the channel D containing the dominant modulations of the target signature *only*. This is meaningful since, in any practical system, independent detection and tracking processes would be used to place a track gate around the predicted target location in each successive frame. The

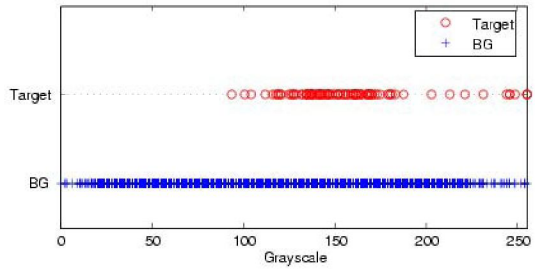

Fig. 2. Scatter plots of pixel gray levels corresponding to target (upper trace) and background (lower trace).

AM function $a_D(\mathbf{m})$ computed according to (4) is shown in Fig. 1(b), while the FM function $\nabla\varphi_D(\mathbf{m})$ obtained using (5) is given by the needle diagram of Fig. 1(c). From this figure, it is clear that the image possesses a rich texture structure well suited for modulation domain representation.

Fig. 2 shows a scatter plot of the pixel gray levels corresponding to the target (upper trace) and background (lower trace) of Fig. 1(a). As is typical in infrared military target detection applications, the raw gray level histograms of the target and background are completely overlapping, a fact which severely limits the utility of the pixel values in discriminating between target and background.

The modulation domain feature space is shown in Fig. 3. Each cross corresponds to a background pixel in Fig. 1(a), while each circle corresponds to a target pixel. In stark contrast to Fig. 2, we see here that the modulation domain representation has clearly succeeded in pulling the target signature out of the background. Indeed, the average nearest neighbor Euclidean distance between target and background pixels in Fig. 2 is identically zero, whereas it is 6.03 units in the feature space of Fig. 3. Although beyond the scope of this paper, it is an important component of our planned future work to evaluate quantitatively the impact of the modulation domain features on discrimination performance in terms of both probability of detection and false alarm rate.

5. CONCLUSION

For the first time, we considered computation of modulation domain features for discriminating targets and backgrounds in FLIR imagery. We outlined a practical computational approach and introduced a new spline-based frequency demodulation algorithm that effectively reduces the approximation errors typical of many previous discrete FM algorithms. While preliminary and qualitative in nature, the examples of Section 4 provide a convincing *proof of concept* clearly demonstrating the power of AM-FM features for enhancing class separability relative to pixel domain features. Our future work will focus on the application of these new features in practical targeting systems and quantitative assessment of the associated processing gain.

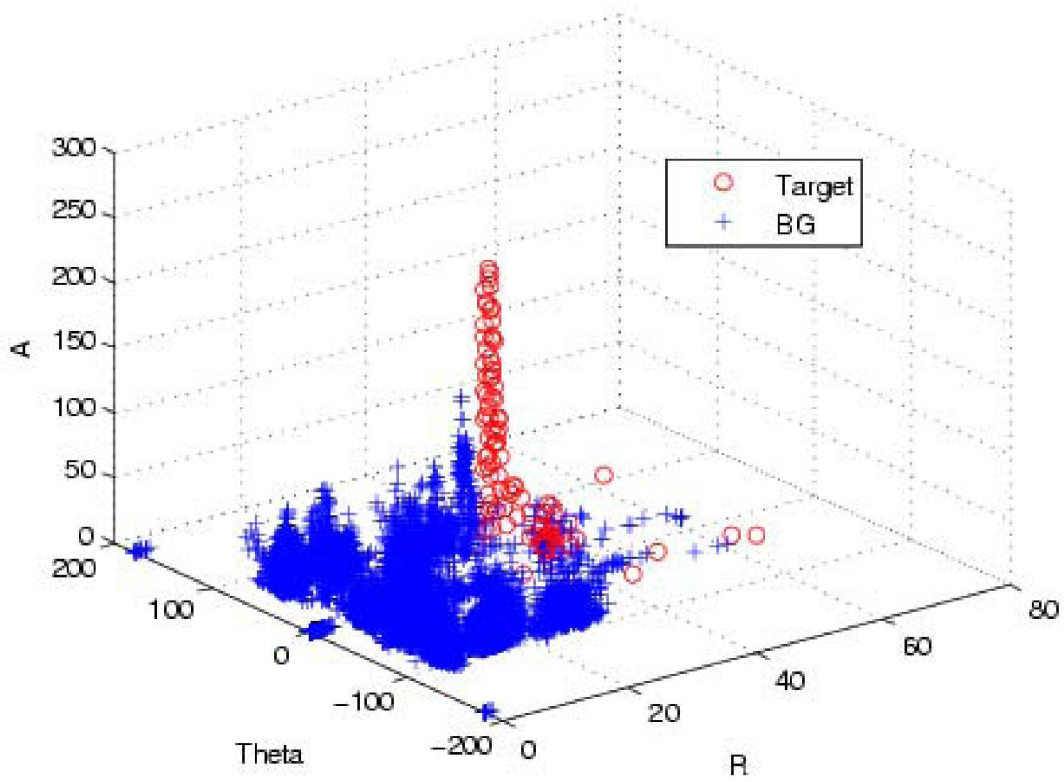

Fig. 3. Scatter plot showing $A(\mathbf{m})$ (pixel units), $R(\mathbf{m})$ (cpi), and $\theta(\mathbf{m})$ (deg) in the modulation domain feature space.

6. REFERENCES

[1] A. Bal and M.S. Alam, "Automatic target tracking in FLIR image sequences using intensity variation function and template modeling," *IEEE Trans. Inst., Measur.*, vol. 54, no. 5, pp. 1846–1852, Oct. 2005.

[2] A. C. Bovik, N. Gopal, T. Emmoth, and A. Restrepo, "Localized measurement of emergent image frequencies by Gabor wavelets," *IEEE Trans. Info. Theory*, vol. 38, no. 2, pp. 691–712, Mar. 1992.

[3] 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. 227–241, Feb. 2000.

[4] P. Maragos and A. C. Bovik, "Image demodulation using multidimensional energy separation," *J. Opt. Soc. Amer. A*, vol. 12, no. 9, pp. 1867–1876, Sep. 1995.

[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.*, Chicago, IL, Oct. 4-7, 1998, pp. 602–606.

[6] 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, 2 edition, 2005.

[7] R. A. Sivley and J. P. Havlicek, "Multidimensional phase unwrapping for consistent APF estimation," in *Proc. IEEE Int'l. Conf. Image Proc.*, Genoa, Italy, Sep. 11-14, 2005, vol. II, pp. 458–461.

[8] D. Dimitriadis and P. Maragos, "An improved energy demodulation using splines," in *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Proc.*, Salt Lake City, UT, May 7-11 2001, vol. 6, pp. 3481–3484.

[9] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing: Part I–theory," *IEEE Trans. Signal Proc.*, vol. 41, no. 2, pp. 821–833, Feb. 1993.

[10] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing: Part II–efficient design and applications," *IEEE Trans. Signal Proc.*, vol. 41, no. 2, pp. 834–848, Feb. 1993.

[11] J. P. Havlicek, A. C. Bovik, and D. Chen, "AM-FM image modeling and Gabor analysis," in *Visual Information Representation, Communication, and Image Processing*, C. W. Chen and Y. Zhang, Eds., pp. 343–385. Marcel Dekker, New York, 1999.