RECONSTRUCTION FROM THE MULTI-COMPONENT AM-FM IMAGE REPRESENTATION

Joseph P. Havlicek, David S. Harding, and Alan C. Bovik

Laboratory for Vision Systems, University of Texas, Austin, TX 78712-1084

ABSTRACT

We present powerful multi-component AM-FM image models capable of efficiently representing complicated nonstationary multi-partite images, and show how the representation can be computed in a practical implementation. With images of this type, important structural and perceptual information is often manifest in the nonstationarities. Highly localized nonlinear operators are used to simultaneously estimate the amplitude and frequency modulating functions associated with each of the multiple components on a pixel-by-pixel basis. For the first time, we also demonstrate image reconstruction from the AM-FM representation.

1. INTRODUCTION

The efficacy of AM-FM modeling techniques for analyzing and characterizing locally coherent nonstationary signals and images has been well established [1–6]. In 1-D, AM-FM models can be efficiently computed using the Teager-Kaiser Energy Operator [4,5]. In terms of analyzing and interpreting the structure, information content, and origin of certain important signals, the inherent ability of AM-FM models to capture local nonstationarities offers significant advantages over traditional time-frequency distributions. Reconstruction of a 1-D signal from the AM-FM model has been demonstrated recently [1].

A multidimensional version of the Teager-Kaiser operator, as well as the demodulation algorithm used in this paper, have been used to compute AM-FM models of images [2,3,6]. In image processing and computational vision, AM-FM modeling has great utility both in estimating instantaneous frequencies on a spatio-spectrally localized basis and in formulating an efficient representation that facilitates analysis. The characterization of images in terms of their spatially localized instantaneous frequency content is fundamental to an increasing variety of multi-dimensional processing techniques including analysis, segmentation and modeling of texture [7,8], shape from texture [9], and emerging techniques in stereopsis [10]. In this paper, we develop the multi-component AM-FM representation for images, outline an algorithm for computing the representation, and for the first time demonstrate reconstruction of images from the computed multi-component AM-FM representation.

This research was supported in part by a grant from the Texas Advanced Research Projects Agency and by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under grant number F49620-93-1-0307.

2. DEMODULATION ALGORITHM

AM-FM modeling is most useful when the images of interest may be accurately represented as a sum of one or more locally coherent complex-valued components, each of the form

$$t(\mathbf{x}) = a(\mathbf{x}) \exp\left[j\varphi(\mathbf{x})\right], \tag{1}$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $t : \mathbb{R}^n \to \mathbb{C}$, $a : \mathbb{R}^n \to [0, \infty)$, and $\varphi : \mathbb{R}^n \to \mathbb{R}$. Such a component can be demodulated using the local nonlinear algorithm

$$a(\mathbf{x}) = |t(\mathbf{x})|, \tag{2}$$

$$\nabla\varphi(\mathbf{x}) = \mathrm{Re}\left[\frac{\nabla t(\mathbf{x})}{jt(\mathbf{x})}\right], \tag{3}$$

which is *exact* for *any* general n-dimensional complex valued AM-FM signal component [3,6].

Complicated multi-partite images, for which there often is no representation in terms of a single component of the form (1) which admits smooth modulating functions, may be better modeled as the real part of the sum

$$t(\mathbf{x}) = \sum_{k=1}^{K} a_k(\mathbf{x}) \exp[j\varphi_k(\mathbf{x})]. \tag{4}$$

In this case, the nonlinear algorithm (2),(3) fails due to interference between the multiple image components. Therefore, prior to demodulation, we separate the individual components using a bank of multiband filters. The filters must be sufficiently spectrally localized to prevent interference between components, but also spatially localized to capture nonstationarities in the locally narrowband components, which may in fact be globally broadband. The design of such a filterbank using a wavelet-like polar tesselation of Gabor functions, which optimally realize the uncertainty principle lower bound on conjoint spatio-spectral localization, can be found elsewhere [3,8,11].

Assuming the filterbank has been properly designed, so that at most one component dominates the response of each channel filter at each pixel, demodulation of the filtered component

$$t_\sigma(\mathbf{x}) = \int_{\mathbb{R}^n} t(\mathbf{x} - \mathbf{p}) g_\sigma(\mathbf{p}) d\mathbf{p}, \tag{5}$$

where $g_\sigma(\mathbf{x})$ is the channel filter with Fourier transform $G_\sigma(\boldsymbol{\Omega})$, can be accomplished using the approximate algo-

0-8186-7310-9/95 $4.00 © 1995 IEEE

rithm

$$\nabla\varphi(\mathbf{x}) \approx \mathrm{Re}\left[\frac{\nabla t_\sigma(\mathbf{x})}{j t_\sigma(\mathbf{x})}\right], \quad (6)$$

$$a(\mathbf{x}) \approx \left|\frac{t_\sigma(\mathbf{x})}{G_\sigma\left(\nabla\varphi(\mathbf{x})\right)}\right|. \quad (7)$$

In deriving (6),(7), we make use of a quasi-eigenfunction approximation [3, 6, 8, 11–13] which tightly bounds the errors in the numerator and denominator of (6) by certain functional norms of $g_\sigma(\mathbf{x})$, $a(\mathbf{x})$, and $\nabla\varphi(\mathbf{x})$. The approximation errors are negligible provided that $g_\sigma(\mathbf{x})$ is well localized spatially and the components of $t(\mathbf{x})$ are reasonably locally coherent. However, the algorithm should not be expected to work well for images which are everywhere highly discontinuous, fractal, or otherwise incoherent.

3. MULTI-COMPONENT REPRESENTATION

The first step in computing the multi-component AM-FM representation of a real image $s(\mathbf{x})$ is to form a complex-valued extension which can be analyzed against (4). We add an imaginary part equal to the 2-D Hilbert transform of $s(\mathbf{x})$, defined by

$$\mathcal{H}[s(\mathbf{x})] = \frac{1}{\pi}\int_{\mathbb{R}} s(\mathbf{x} - y\mathbf{e}_i)\frac{dy}{y} = s(\mathbf{x}) * \frac{\delta(\mathbf{x}^T\mathbf{e}_j)}{\pi\mathbf{x}^T\mathbf{e}_i}, \quad (8)$$

where $\mathbf{e}_i = [1,0]^T$, $\mathbf{e}_j = [0,1]^T$, δ is the Dirac delta, and the integral is interpreted as a Cauchy principle value. We refer to the extension $t(\mathbf{x}) = s(\mathbf{x}) + j\mathcal{H}[s(\mathbf{x})]$ as the *analytic image associated with* $s(\mathbf{x})$. The spectrum $T(\mathbf{\Omega})$, which is supported only in quadrants I & IV of the frequency plain, results from removing all spectral redundancy from the conjugate symmetric function $S(\mathbf{\Omega})$.

The filterbank is invoked on the analytic image, and the filtered demodulation algorithm (6),(7) is applied to all channel responses. Hence, each channel produces observations $a(\mathbf{x})$ and $\nabla\varphi(\mathbf{x})$ at every pixel. We use a track processor based on computationally efficient 2nd-order 1-D Kalman filters to track the individual components across the channels on a pixel-by-pixel basis [3]. A block diagram of this scheme is shown in Figure 1.

4. RECONSTRUCTION

Reconstructing an image component from the discrete AM-FM representation $a(\mathbf{k})$, $\nabla\varphi(\mathbf{k})$ is an overdetermined problem. Given as an initial condition one complex valued sample obtained from the channel filter used to estimate the component's instantaneous amplitude and frequency, the instantaneous phase $\varphi(\mathbf{k})$ can be reconstructed along any contiguous path of pixels by simply summing the frequency estimates. The component can then be reconstructed by substituting the amplitude estimates and reconstruction of $\varphi(\mathbf{k})$ directly into the model (1).

In practice, however, the frequency estimates will contain errors which arise from the approximation inherent in the algorithm (6), (7), as well as from interference by other components and out-of-band noise. If only a single sample of the component were used to determine the initial value of the instantaneous phase, the deleterious effects of these errors on the reconstructed component would be cumulative.

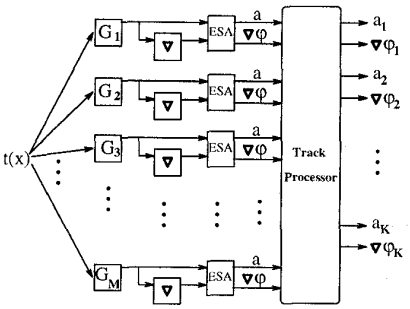

Figure 1: System for computing the multi-component AM-FM representation.

While sophisticated interpolation techniques could be used to compensate for these errors, we employ a simple but effective two-part approach for making the reconstruction algorithm more robust. First, instead of saving only a single sample of the estimated component, we save samples on a rectangular grid, and reconstruct the component on each square of the grid independently. We use the estimated component sample at the upper left corner of the square to determine the initial value of the instantaneous phase. Second, we compute two estimates of the phase at each pixel: one using the phase of a horizontally neighboring pixel, and one using the phase of a vertically neighboring pixel. The component is then reconstructed using the average of these two estimates. Finally, the multi-partite image $s(\mathbf{x})$ is reconstructed by summing the individually reconstructed components, according to the model (4).

5. EXAMPLES

In Figure 2, we compute the multi-component AM-FM representation of a synthetically generated two-component image exhibiting significant 2-D nonstationarity, and then reconstruct. The original image is shown in the (a) part of the figure. True values for components $C1$ and $C2$ are shown in Figures 2 (b) and (c), respectively. The presence of two components was correctly identified by the track processor. The amplitude estimates for $C1$ are shown in Figure 2 (d), while the horizontal and vertical frequency estimates are given in the (e) and (f) parts of the figure, respectively. For $C2$, the amplitude, horizontal frequency, and vertical frequency estimates are given in the (g), (h), and (i) parts of the figure. To avoid edge effects from the channel filters, tracking and reconstruction were not performed on the outside 16 rows and columns of the image. Note how smooth the AM-FM representation is, despite the fact that there are rapid variations in the image. Significant compression of the estimated quantities could be achieved, *e.g.* through

simple linear predictive coding. The estimated quantities are in near perfect agreement with the true values, except for a small region of oscillatory behavior in $C1$ where the vertical frequency approaches DC (which is not surprising, given that the concept of instantaneous frequency is less meaningful near DC).

The individual components reconstructed using a 32×32 pixel grid of initial conditions are shown in the (k) and (l) parts of the figure. The reconstructed components, which are virtually indistinguishable from the true values visually, were summed to obtain the reconstructed image shown in Figure 2 (j).

Figure 3 (a) shows the complicated, nonstationary multipartite natural Brodatz texture image *tree*. In general, AM-FM analysis of such images is challenging, due primarily to the presence of an unknown number of components which may each be supported only on irregularly shaped subimages. This situation presents significant difficulties for our current simplistic 1-D track processor, which depends on components being supported over the entire tracked region. Furthermore, the presence of many components and harmonics closely spaced in frequency gives rise to errors in the filtered demodulation algorithm due to cross-component interference.

Nevertheless, we were able to obtain the excellent reconstruction shown in Figure 3 (d) by performing track processing independently over several subregions of the image, with the side effect that some blocking artifacts are visible. For this image, the track processor was not able to automatically detect the number of components. Many, many tracks were generated in each subregion. The reconstruction in Figure 3 (d) comprises 41 hand-selected components, five of which are shown juxtaposed in the (c) part of the figure.

Reconstruction was performed on a 4×4 pixel grid of estimated component samples. The low-frequency component shown in Figure 3 (b) was extracted by linear filtering prior to computation of the AM-FM representation and added back in to the reconstruction. This component is not ideally suited for AM-FM modelling, because at very low spatial frequencies it becomes ambiguous which features should be interpreted as AM as opposed to FM. After incorporating the low-frequency component, however, we still have a smooth representation amenable to analysis and compression. Most of the information present in the image which seems to be missing from the reconstruction is high-frequency and incoherent in nature.

6. FUTURE WORK

Important future work remaining in this area includes developing methods for overcoming the extremely difficult problems in treating complicated natural images. In particular, this will involve developing an approach for automatically detecting the number of AM-FM components that are present and for effectively tracking components over irregularly shaped regions of support.

7. REFERENCES

[1] J. P. Havlicek, D. S. Harding, and A. C. Bovik, "Multi-component signal demodulation and reconstruction using AM-FM modulation models", in *Proc. 1995 IEEE Workshop Nonlin. Signal and Image Proc.*, Neos Marmaras, Halkidiki, Greece, June 20 - 22 1995.

[2] P. Maragos and A. C. Bovik, "Demodulation of images modeled by amplitude-frequency modulation using multidimensional energy separation", in *Proc. IEEE Int'l. Conf. Image Proc.*, Austin, TX, 1994, pp. III421–III425.

[3] J. P. Havlicek and A. C. Bovik, "Multi-component AM-FM image models and wavelet-based demodulation with component tracking", in *Proc. IEEE Int'l. Conf. Image Proc.*, Austin, TX, November 13-16 1994, pp. I41 – I45.

[4] P. Maragos, J. F. Kaiser, and T. F. Quatieri, "Energy separation in signal modulations with applications to speech analysis", *IEEE. Trans. Signal Proc.*, vol. SP-41, no. 10, pp. 3024–3051, October 1993.

[5] A. C. Bovik, P. Maragos, and T. F. Quatieri, "AM-FM energy detection and separation in noise using multiband energy operators", *IEEE. Trans. Signal Proc.*, vol. SP-41, no. 12, pp. 3245–3265, December 1993.

[6] J. P. Havlicek, A. C. Bovik, and P. Maragos, "Modulation models for image processing and wavelet-based image demodulation", in *Proc. 26th IEEE Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, October 26 - 28 1992, pp. 805–810.

[7] A. C. Bovik, M. Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters", *IEEE. Trans. Pattern Anal. Machine Intell.*, vol. PAMI-12, no. 1, pp. 55–73, January 1990.

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

[9] B. J. Super and A. C. Bovik, "Solution to shape-from-texture by wavelet-based measurement of local spectral moments", Tech. Rept. TR-92-5-80, Computer and Vision Research Center, The University of Texas at Austin, November 1991.

[10] T. Y. Chen, A. C. Bovik, and B. J. Super, "Multiscale stereopsis via gabor filter phase response", in *Proc. IEEE Int'l. Conf. Syst., Man, and Cyber.*, SanAntonio, TX, October 2 - 5 1994, pp. 55 – 60.

[11] J. P. Havlicek and A. C. Bovik, "AM-FM models, the analytic image, and nonlinear demodulation techniques", Tech. Rept. TR-95-001, Center for Vision and Image Sciences, The University of Texas at Austin, March 1995.

[12] A. C. Bovik, "A bound involving n-dimensional instantaneous frequency", *IEEE. Trans. Circ. Syst.*, vol. CAS-38, no. 11, pp. 1389–1390, November 1991.

[13] A. C. Bovik, J. P. Havlicek, and M. D. Desai, "Theorems for discrete filtered modulated signals", in *Proc. IEEE Int'l. Conf. Acoust., Speech, Signal Proc.*, Minneapolis, MN, April 27 - 30 1993, pp. III153–III156.

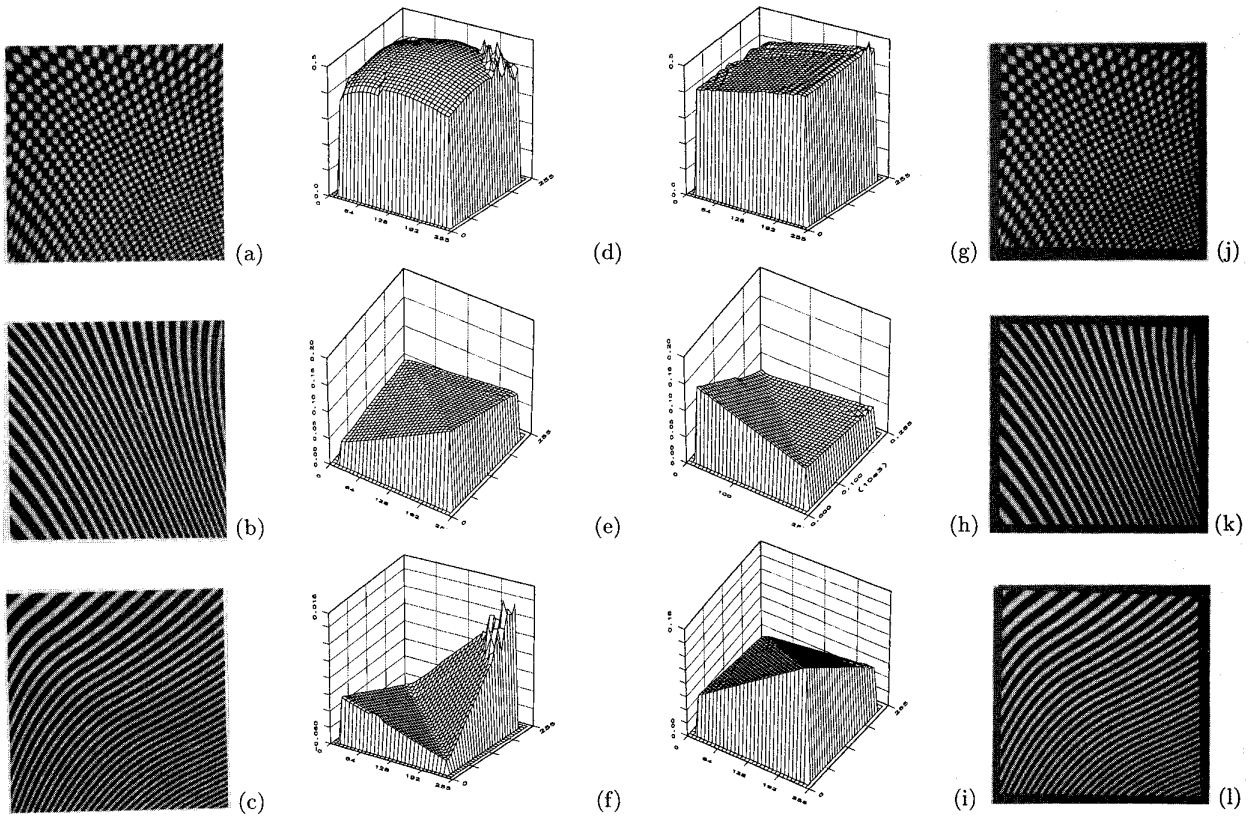

Figure 2: Multi-component AM-FM representation and reconstruction of a synthetic image. (a) Nonstationary two-component synthetic image. (b) True values for component one. (c) True values for component two. (d) Amplitude estimates for component one. (e) Horizontal frequency estimates for component one. (f) Vertical frequency estimates for component one. (g) Amplitude estimates for component two. (h) Horizontal frequency estimates for component two. (i) Vertical frequency estimates for component two. (j) Reconstructed image. (k) Reconstruction of component one. (l) Reconstruction of component two.

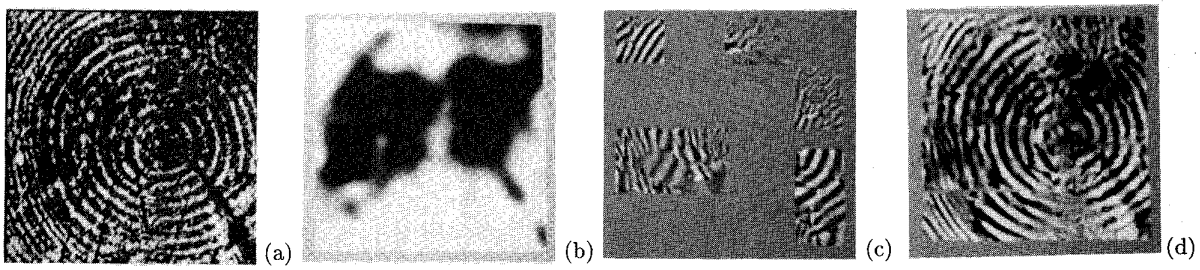

Figure 3: Reconstruction of a complicated natural multi-partite texture image. (a) Tree image. (b) Low-pass component extracted by linear filtering. (c) Five of the 41 components used to reconstruct. (d) Reconstruction from 41 hand-selected AM-FM components.