First Results in Perceptually-Based AM-FM Image Filtering

Chuong T. Nguyen,[†] Roy A. Sivley,[‡] and Joseph P. Havlicek[†]
[†]School of Electrical and Computer Engineering, University of Oklahoma, Norman, OK, USA
[‡]BAE Systems, Burlington, MA, USA

Abstract

We combine an adaptation of the steerable image pyramid subband decomposition with a spline-based perfect reconstruction demodulation algorithm to obtain an invertible AM-FM image transform. For the first time, we achieve perceptually-based signal processing goals by applying filtering operations directly to the computed subband amplitude and frequency modulations. The results are dramatic and would be difficult or impossible to obtain by linear processing. In our most interesting example, a simple AM-FM filter succeeds in smoothly and naturally removing the bands from the hat in the well-known Lena image.

1. Introduction

The Fourier series and integral represent a signal in terms of a sum of complex-valued sinusoids each having an amplitude and frequency that are everywhere constant. We refer to the Fourier components as *stationary* in this sense. Linear translation invariant (LTI) filters perform signal processing by multiplying each input Fourier component by a complex eigenvalue. For a real impulse response, the effects are twofold: for each component, the amplitude is scaled by the magnitude of the eigenvalue and the phase is shifted by the angle of the eigenvalue. The collection of eigenvalues is called the frequency response and is given by the Fourier transform of the impulse response. In designing an LTI filter, one specifies the frequency response to achieve a desired signal processing goal.

In many applications, however, the desired signal processing goal is based on human perception of the input and output signals. In such cases, it may not be straightforward to express the goal directly in terms of stationary Fourier amplitudes and frequencies. For example, consider the *Trio* from Bach's first Brandenburg Concerto. There are two oboes and a bassoon. We hear this piece as a superposition of the three musical lines played by the individual instruments. Each line is a sequence of notes characterized by a fundamental frequency, a weighted overtone series, and an amplitude envelope. Thus, our aural perception is more naturally related to *nonstationary* notions of amplitude and frequency than to any stationary interpretation in terms of sinusoids with amplitudes and frequencies that remain constant throughout the entire movement.

There is strong psychophysical and physiological evidence that biological vision systems also respond directly to spatially and spectrally localized bundles of amplitudes and frequencies [4, 5, 10, 11]. The idea of analyzing images in terms of nonstationary amplitude and frequency was introduced in [1, 2] and developed in [6, 7, 12], resulting in the multicomponent AM-FM model

$$t(\mathbf{x}) = \sum_{m=1}^{M} a_m(\mathbf{x}) \cos[\varphi_m(\mathbf{x})] \tag{1}$$

for an image $t\colon \mathbb{R}^2 \to \mathbb{R}$. In (1), the image is interpreted as a sum of M nonstationary AM-FM components

$$t_m(\mathbf{x}) = a_m(\mathbf{x}) \cos[\varphi(\mathbf{x})]. \tag{2}$$

Intuitively, the AM function a_m characterizes the local contrast or envelope of t_m, while the FM function $\nabla\varphi_m$ characterizes the local orientation and granularity of the visible patterns.

Estimation of the AM and FM functions from the image is a problem that is difficult and ill-posed on several levels [9]. First, there is the question of how many components there should be and how they can be obtained from $t(\mathbf{x})$. Even if the individual components t_m could be observed directly, the demodulation problem is ill-posed in the sense that, for any given t_m, there are an uncountable infinity of functions a_m and $\nabla\varphi_m$ satisfying (2). Biologically motivated filterbanks of one type or another have typically been used to isolate the components, but closed form expressions for the AM and FM functions in terms of the filter responses cannot usually be obtained. Demodulation algorithms such as those given in [7, 12, 14] are based on approximations yielding estimates for the AM and FM functions that are demonstrably useful in a variety of applications but that generally preclude the possibility of reconstructing the image t from the computed AM-FM model.

We first proposed the idea of a perfect reconstruction AM-FM image model for performing modulation domain signal processing in [9, 17, 18]. This concept is illustrated in the block diagram of Fig. 1. Although perfect reconstruction was achieved in [9, 17, 18] with the *Signal Processing* block of Fig. 1 set to the "do nothing" operation, we did not succeed in demonstrating any perceptually relevant modulation domain image processing. The main reason is that we were unable to design a perfect reconstruction filterbank capable of delivering components t_m with sufficient local coherency [2,7] to provide good correspondence between the computed AM and FM functions and visual perception of the salient image structure.

In this paper, we replace the nonseparable wavelet filterbank used in [9, 17] with an adaptation of the steerable pyramid decomposition given in [15, 16] to obtain a model (1) in terms of perceptually relevant, locally coherent components (2). We use this model to perform perceptually motivated image processing in the

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.

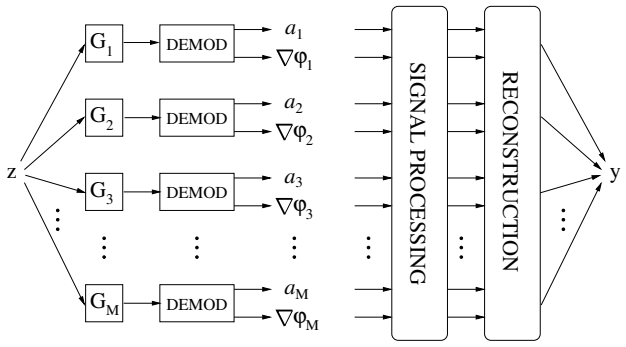

Figure 1. System block diagram for performing AM-FM image processing.

modulation domain for the first time. A striking example is given in Fig. 2(h), where we have used AM-FM filtering to erase the bands from Lena's hat – a result that would be difficult or impossible to obtain with LTI processing.

2. Invertible AM-FM Image Model

In this section, we discuss the techniques we have used to realize the block diagram of Fig. 1 subject to the following: 1) the signal processing operation is set to the "do nothing" operation, 2) the AM-FM transform is invertible in the sense that the input and output images look the same, and 3) the AM and FM functions are relevant to visual perception of the image.

2.1. Demodulation

As we mentioned in Section 1, the problem of associating a pair of modulating functions a_m and $\nabla\varphi_m$ to a real-valued image component (2) is ill-posed because the solution is not unique. By contrast, the AM and FM functions of any complex-valued component $z_m\colon \mathbb{R}^2\to\mathbb{C}$ are unique up to equivalence classes of functions that agree a.e. with respect to Lebesgue measure. Therefore, to disambiguate the problem of demodulating the real-valued component (2), we construct a complex extension

$$z_m(\mathbf{x}) = t_m(\mathbf{x}) + jq_m(\mathbf{x}) = a_m(\mathbf{x})\exp[j\varphi_m(\mathbf{x})], \quad (3)$$

where, with $\mathbf{x} = [x_1\ x_2]^T$, $q_m(\mathbf{x})$ is given by the partial Hilbert transform [8]

$$q_m(\mathbf{x}) = \text{p.v.}\frac{1}{\pi}\int_{\mathbb{R}} t_m(x_1-\xi, x_2)\frac{d\xi}{\xi}. \quad (4)$$

With this approach, the complex analytic image z_m admits many but not all of the attractive properties of the 1-D analytic signal.[1] The AM and FM functions of the complex-valued component (3) may be obtained by [7]

$$\nabla\varphi_m(\mathbf{x}) = \frac{\nabla z_m(\mathbf{x})}{jz_m(\mathbf{x})}, \quad (5)$$

$$a_m(\mathbf{x}) = |z_m(\mathbf{x})|. \quad (6)$$

[1] In general, the analytic image $z_m(\mathbf{x})$ fails to satisfy the multidimensional Cauchy-Riemann equations.

Unfortunately, approximations must be applied to discretize (5), and hence a corresponding *exact* discrete demodulation algorithm is not known. Here, we adopt the solution that was given in [17, 18]. We apply the approximate discretization of (5) given in [7] to obtain an initial estimate of the frequency field. This estimate is used to perform 2-D phase unwrapping on the complex component (3). We fit the unwrapped phase with a cubic tensor product spline and differentiate analytically to obtain $\nabla\varphi_m(\mathbf{x})$. The AM function is calculated according to (6). Since the Hilbert transform (4) is linear, it may be applied directly to $t(\mathbf{x})$ in (1) to at once generate all M complex components (3) according to

$$z(\mathbf{x}) = t(\mathbf{x}) + jq(\mathbf{x}) = \sum_{m=1}^{M} t_m(\mathbf{x}) + jq_m(\mathbf{x}). \quad (7)$$

2.2. Decomposition into Components

A major difficulty with obtaining perfect reconstruction (or approximately perfect reconstruction) from the AM-FM models given in [7, 13] is that the Gabor filterbanks used to isolate components in those approaches fail to admit compact frequency support. Hence, Gabor filters cannot be used to formulate a perfect reconstruction filterbank. Nevertheless, Gabor filters admit many attractive properties. They are C^∞ and they uniquely realize the Heisenberg-Weyl inequality limit on joint time-frequency localization in $L^2(\mathbb{R}^2)$ [5]. Gabor filterbanks tend to deliver locally coherent components that correspond well to human visual perception. The obvious question is: *how can one construct Gabor-like filterbanks that retain the desirable properties but also provide approximate perfect reconstruction?*

In [9,17–19], we constructed cartesian-separable perfect reconstruction QMF filterbanks with good joint localization properties. As a consequence of separability, each filterbank channel was supported in all four quadrants of the 2-D frequency plane. Hence the resulting components (3) contained multiple orientations and were not locally coherent [3, 20]. However, we showed that each separable channel admits loci of zeros in the frequency plane that can be utilized to divide it into a number of *nonseparable* subchannels that are orientation selective. Nevertheless, three main difficulties prevented this approach from working well in practice. First, the required number of nonseparable subchannels is large, on the order of 500 for a 256^2 image. Second, there are always some channels that are low pass in one dimension only and hence contain multiple approximately orthogonal orientations that cannot be further decomposed. Third, the interband aliasing that occurs in the synthesis filterbank makes it difficult or impossible to reliably predict the effects of modulation domain signal processing operations performed between the analysis and synthesis banks.

Here, these problems are overcome by adapting the steerable pyramid decomposition developed in [15, 16]. The steerable pyramid decomposes an image into scale and orientation selective subbands based on polar-separable directional derivative operators. Because the analysis channels are self inverting, interband aliasing is effectively eliminated in the synthesis filterbank. Although the steerable pyramid does not provide perfect reconstruction, the reconstruction errors are typically incoherent and small in L^2 norm, and thus visually insignificant.

We implemented the steerable pyramid as described in [15, 16] with four radial frequency channels at each of $k = 16$ orientations for a total of 64 subbands in the recursive subsystem of the decomposition. Since the high pass channel (denoted $H_0(\omega)$ in [15]) is

not orientation selective, we further decomposed it into 16 orientation selective subchannels. In the notation of [15], the frequency responses of these subchannels are given by $B_i(\boldsymbol{\omega}) = A(\theta-\theta_i) = \cos^{15}(\theta-\theta_i)$, $0 \le i \le 15$. The frequency response of the $i = 4$ subchannel is shown in Fig. 2(a). The adapted steerable pyramid thus consists of 81 subbands including 16 orientation selective high pass channels, 64 orientation and radially selective recursive channels, and one low pass residual. Demodulation, AM-FM filtering, and AM-FM reconstruction were performed between the analysis and synthesis filterbanks.

3. Perceptually-Based AM-FM Image Filtering

In this section, we demonstrate *for the first time* nontrivial AM-FM image filtering including both analysis and synthesis of the signal processing result. The AM-FM image model $\{a_m, \nabla\varphi_m\}_{m\in[1,81]}$ was computed between the analysis and synthesis filterbanks as described in Section 2.1. Filtering was applied to the AM and FM functions to achieve perceptually-based signal processing goals. AM-FM reconstruction was performed using the filtered modulating functions. The resulting processed subband image components were then input to the reconstruction filterbank to obtain the signal processed image.

In discussing the experiments, we denote the filtered AM and FM functions by $\{\widehat{a}_m, \nabla\widehat{\varphi}_m\}_{m\in[1,81]}$. The unfiltered FM functions are given in polar form by $r_m = |\nabla\varphi_m|$ and $\psi_m = \arg\nabla\varphi_m$. Likewise, in polar form the filtered FM functions are given by $\widehat{r}_m = |\nabla\widehat{\varphi}_m|$ and $\widehat{\psi}_m = \arg\nabla\widehat{\varphi}_m$.

Example 1: Orientation Selective Attenuation of Structure in a Synthetic Image. Fig. 2(b) shows an isotropic synthetic radial chirp image where the phase is a quadratic function of the distance from the center of the image. The signal processing goal was to perform orientation selective attenuation along the main diagonals where the orientation is an odd multiple of $\pi/4$. With $\delta_m(\mathbf{x}) = |(|\psi_m(\mathbf{x})| - \pi/4)|$, the AM-FM filtering operation is given by

$$\widehat{a}_m(\mathbf{x}) = \begin{cases} 16\delta_m(\mathbf{x})a_m(\mathbf{x})/\pi, & \delta_m(\mathbf{x}) < \pi/16, \\ a_m(\mathbf{x}), & \text{otherwise} \end{cases} \tag{8}$$

and $\nabla\widehat{\varphi}_m(\mathbf{x}) = \nabla\varphi_m(\mathbf{x})$. The modulation domain notch filter (8) may be intuitively understood as follows: for subband pixels $z_m(\mathbf{x})$ with an orientation $\psi_m(\mathbf{x})$ that differs from an odd multiple of $\pi/4$ by $\pi/16$ or less, the AM function $a_m(\mathbf{x})$ is multiplied by a gain that is linear in $\delta_m(\mathbf{x})$. Thus, the amplitude is zeroed if $\delta_m(\mathbf{x}) = 0$ and is multiplied by one if $\delta_m(\mathbf{x}) \ge \pi/16$. The result is shown in Fig. 2(c), where it is clear that the perceptually-based signal processing goal has been achieved.

Example 2: Orientation Selective Attenuation of Structure in a Natural Image. The natural wood grain texture image *Tree* given in Fig. 2(d) bears strong similarity to the synthetic chirp in Fig. 2(b). The experiment of Example 1 was repeated on this image, but with a more aggressive notch of half-width $\pi/8$ in order to attenuate a wider band of orientations. The result is given in Fig. 2(e) where it may again be seen that the perceptually-based signal processing goal has been achieved. The effect of this AM-FM filtering is clearly visible in the left upper and lower quadrants of the image. Interesting subtle effects are also present. For example, careful examination of the center of the original image in Fig. 2(d), just to the right of the knot of the wood grain, reveals a small "hook" that is oriented along the main diagonal. Consistent with the processing goal, this hook is smoothly but totally obliterated in the result image of Fig. 2(e).

Example 3: Isotropic AM-FM Contrast Sharpening Based on Frequency Magnitude. Using again the *Tree* image of Fig. 2(d), the signal processing goal in this case was to perform contrast enhancement and sharpening by selectively amplifying the AM functions a_m for a high pass band of magnitude frequencies r_m. With $T = 2$ rad/pix, the AM-FM filtering operation is given by

$$\widehat{a}_m(\mathbf{x}) = \begin{cases} a_m(\mathbf{x})\exp[r_m(\mathbf{x})/2], & r_m(\mathbf{x}) > T, \\ a_m(\mathbf{x}), & \text{otherwise} \end{cases} \tag{9}$$

and $\nabla\widehat{\varphi}_m(\mathbf{x}) = \nabla\varphi_m(\mathbf{x})$. The result is shown in Fig. 2(f) where it is clear that the perceptually-based signal processing goal has once again been achieved.

Example 4: Spatially Selective Removal of Oriented Structure from a Natural Image. The signal processing goal was to remove the bands from Lena's hat. The original image is given in Fig. 2(g). Let $\mathcal{X}$ denote the interior of the black rectangle shown superimposed on Fig. 2(g). AM-FM filtering was applied to the components $z_m(\mathbf{x})$, but only for $\mathbf{x} \in \mathcal{X}$. With $\delta_m(\mathbf{x}) = |\psi_m(\mathbf{x}) - \pi/4|$, the AM-FM filtering operation is given by

$$\widehat{a}_m(\mathbf{x}) = \begin{cases} 0, & \delta_m(\mathbf{x}) < \pi/8 \text{ and } \mathbf{x} \in \mathcal{X}, \\ a_m(\mathbf{x}), & \text{otherwise} \end{cases} \tag{10}$$

and $\nabla\widehat{\varphi}_m(\mathbf{x}) = \nabla\varphi_m(\mathbf{x})$. As shown in Fig. 2(h), the perceptually-based signal processing goal was achieved with a smooth, natural appearance. There are a few unwanted artifacts that result from the fact that we were quite imprecise in our specification the spatial region $\mathcal{X}$ desired for processing. For example, the central portions of the upper and lower edges of the hat brim were attenuated, as were certain orientations in the upper portion of the feather, and a slight shadow was induced on Lena's forehead. Interestingly, the shadow appearing on the upper portion of the hat was virtually unaffected. All of the unwanted artifacts could be avoided by specifying the region $\mathcal{X}$ more precisely, but doing so would require increased effort in designing the filter.

4. Conclusion

For the first time, we demonstrated AM-FM image filtering including an analysis transform, signal processing in the modulation domain, and a synthesis transform. We designed simple AM-FM filters that were highly effective in achieving perceptually-based signal processing goals that would be difficult or impossible to obtain with LTI filters. Important remaining open problems include the investigation of more sophisticated AM-FM filter algorithms and amelioration of the singularity in the partial Hilbert transform which perturbs the frequency estimates lying near the vertical axis of the 2-D Fourier plane.

References

[1] A. C. Bovik, M. Clark, and W. S. Geisler. Multichannel texture analysis using localized spatial filters. *IEEE Trans. Pattern Anal. Machine Intell.*, 12(1):55–73, Jan. 1990.

[2] A. C. Bovik, N. Gopal, T. Emmoth, and A. Restrepo. Localized measurement of emergent image frequencies by Gabor wavelets. *IEEE Trans. Info. Theory*, 38(2):691–712, Mar. 1992.

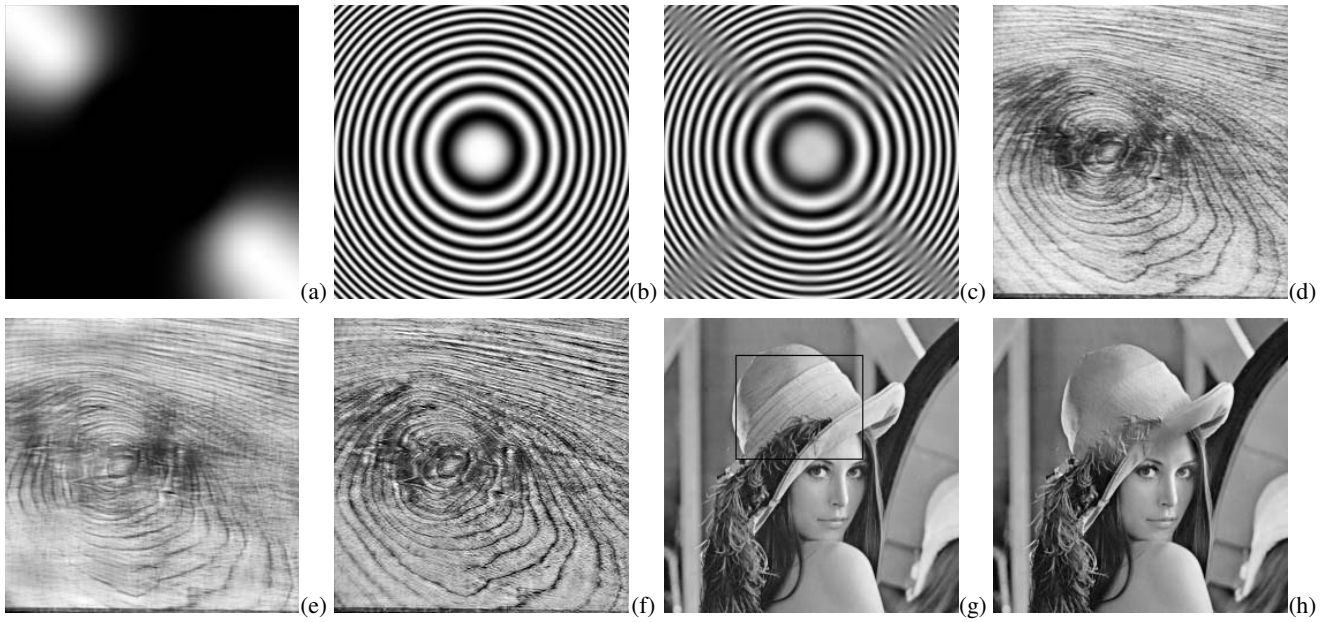

Figure 2. Examples. (a) Hi pass orientation selective subchannel obtained by angularly decomposing the standard steerable pyramid high pass channel. (b) *Radial chirp* **image. (c) Result of AM-FM notch filter centered at odd multiples of $\pi/4$. (d)** *Tree* **image. (e) Result of AM-FM notch filter centered at odd multiples of $\pi/4$; the notch is twice as wide as in (c). (f) Frequency selective AM-FM contrast enhancement. (g)** *Lena* **image. Spatially selective processing was applied only within the black rectangle. (h) AM-FM filtering result. The perceptually-based signal processing goal of removing the bands from Lena's hat was achieved.**

[3] L. Cohen. *Time-Frequency Analysis*. Prentice Hall, Englewood Cliffs, NJ, 1995.

[4] L. K. Cormack. Computational models of early human vision. In A. C. Bovik, editor, *Handbook of Image and Video Processing*, pages 325–345. Elsevier Academic Press, Burlington, MA, 2nd edition, 2005.

[5] 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(7):1160–1169, Jul. 1985.

[6] J. P. Havlicek, D. S. Harding, and A. C. Bovik. The multi-component AM-FM image representation. *IEEE Trans. Image Process.*, 5(6):1094–1100, Jun. 1996.

[7] J. P. Havlicek, D. S. Harding, and A. C. Bovik. Multi-dimensional quasi-eigenfunction approximations and multi-component AM-FM models. *IEEE Trans. Image Process.*, 9(2):227–242, Feb. 2000.

[8] J. P. Havlicek, J. W. Havlicek, and A. C. Bovik. The analytic image. In *Proc. IEEE Int'l. Conf. Image Process.*, Santa Barbara, CA, Oct. 26-29, 1997.

[9] J. P. Havlicek, P. C. Tay, and A. C. Bovik. AM-FM image models: Fundamental techniques and emerging trends. In A. C. Bovik, editor, *Handbook of Image and Video Processing*, pages 377–395. Elsevier Academic Press, Burlington, MA, 2nd edition, 2005.

[10] J. P. Jones, A. Stepnoski, and L. A. Palmer. The two-dimensional spectral structure of simple receptive fields in cat striate cortex. *J. Neurophysiol.*, 58(6):1212–1232, 1987.

[11] J. Malik and P. Perona. Preattentive texture discrimination with early vision mechanisms. *J. Opt. Soc. Am. A*, 7(5):923–932, May 1990.

[12] P. Maragos and A. C. Bovik. Image demodulation using multidimensional energy separation. *J. Opt. Soc. Amer. A*, 12(9):1867–1876, Sep. 1995.

[13] P. Maragos, J. F. Kaiser, and T. F. Quatieri. On amplitude and frequency demodulation using energy operators. *IEEE Trans. Signal Process.*, 41(4):1532–1550, Apr. 1993.

[14] M. Pattichis and A. Bovik. Analyzing image structure by multidimensional frequency modulation. *IEEE Trans. Pattern Anal., Machine Intel.*, 29(5):753–766, May 2007.

[15] E. P. Simoncelli and W. T. Freeman. The steerable pyramid: a flexible architecture for multi-scale derivative computation. In *Proc. IEEE Int'l. Conf. Image Process.*, pages 444–447, Washington, DC., Oct. 23-26, 1995.

[16] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger. Shiftable multi-scale transform. *IEEE Trans. Info. Theory*, 38(2):587–607, March. 1992.

[17] R. A. Sivley and J. P. Havlicek. Perfect reconstruction AM-FM image models. In *Proc. IEEE Int'l. Conf. Image Process.*, pages 2125–2128, Atlanta, GA, Oct. 8-11, 2006.

[18] R. A. Sivley and J. P. Havlicek. A spline-based framework for perfrect reconstruction AM-FM models. In *Proc. IEEE Southwest Symp. Image Anal., Interp.*, pages 198–202, Denver, CO, Mar. 26-28 2006.

[19] P. C. Tay and J. P. Havlicek. Joint uncertainty measures for maximally decimated M-channel prime factor cascaded wavelet filter banks. In *Proc. IEEE Int'l. Conf. Image Process.*, pages 1033–1036, Barcelona, Spain, Sep. 14-17, 2003, vol. I.

[20] D. Wei and A. C. Bovik. On the instantaneous frequencies of multicomponent AM-FM signals. *IEEE Signal Process. Let.*, 5(4):84–86, Apr. 1998.