ENERGY DEMODULATION OF TWO–COMPONENT AM–FM SIGNALS WITH APPLICATION TO SPEAKER SEPARATION

Balasubramaniam Santhanam and Petros Maragos

School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332 - 0250, USA
santhan@ece.gatech.edu

ABSTRACT

In this paper, an efficient low-complexity algorithm is presented for the separation and demodulation of two–component AM–FM signals and applied to the separation of voice–modulated FM signals. The proposed algorithm is based on the generating differential equation of the mixture signal and nonlinear differential energy operators.

1. INTRODUCTION

Monocomponent AM–FM signals of the form $x(t) = a(t)\cos\left(\int_0^t \omega(\tau)d\tau\right)$ are very useful in analog communication systems for the transmission of information over communication channels. Recently they have also found applications in speech processing, where they are used to model speech resonances. To demodulate $x(t)$ into its amplitude envelope $|a(t)|$ and instantaneous frequency signal $\omega(t)$ the *energy separation algorithm* (ESA) was recently proposed in [1, 2]:

$$\sqrt{\frac{\Psi(\dot{x})}{\Psi(x)}} \approx \omega(t) \quad , \quad \frac{\Psi(x)}{\sqrt{\Psi(\dot{x})}} \approx |a(t)|,$$

where $\Psi(x) \quad \stackrel{\triangle}{=} \quad (\dot{x})^2 - x\ddot{x}$ is the Teager–Kaiser energy operator [3] and dots denote time derivatives. Two–component AM–FM signals of the form

$$x(t) = a_1(t)\cos\left(\int_0^t \omega_1(\tau)d\tau\right) + a_2(t)\cos\left(\int_0^t \omega_2(\tau)d\tau\right)$$

are also of particular interest and have been used to model co–channel and adjacent interferences over communication channels. Separation and demodulation of these signals can be accomplished via the use of several methods:
1) multiband–ESA (MESA) [4]. 2) LMS algorithm [5]. 3) Toeplitz operators (ITD) algorithm [6]. 4) *Hankel rank reduction* (HRR) algorithm [7]. 5) CC–DPLL algorithm [8].

In this paper, we present the *energy demodulation of mixtures* (EDM) algorithm, that is based on higher-order energy operators [9], for the separation and demodulation of two–component AM–FM signals. We also present its application to the separation of two–component voice–modulated FM signals.

This research work was supported by the National Science Foundation under Grant MIP-93963091.

2. CONTINUOUS–TIME ALGORITHM

The two–component AM–FM signal is modeled instantaneously as a two–component sinusoidal signal of the form

$$x(t) = a_1\cos(\omega_1 t + \theta_1) + a_2\cos(\omega_2 t + \theta_2).$$

This mixture signal satisfies the following fourth-order *generating differential equation* (GDE):

$$x^{(4)} + c_1\ddot{x} + c_2 x = 0,$$

where $x^{(n)} = d^n x/dt^n$ and

$$c_1 = (\omega_1^2 + \omega_2^2) \quad , \quad c_2 = \omega_1^2\omega_2^2.$$

The k^{th}-order differential energy operator is defined as [9]

$$\Upsilon_k(x) \quad \stackrel{\triangle}{=} \quad \dot{x}x^{(k-1)} - xx^{(k)}.$$

These operators measure the higher-order energies of a classical mono–component harmonic oscillator, which has been subjected to a displacement of $x(t)$, normalized to half unit mass [9]. As a special case for $k = 2$, we obtain the Teager-Kaiser energy operator $\Upsilon_2 \equiv \Psi$. Using the GDE and its derivative and solving the resultant 2×2 linear system of equations yields the following expressions for the coefficients:

$$c_1 = -\frac{\Upsilon_5(x)}{\Upsilon_3(x)} \quad , \quad c_2 = \frac{\Upsilon_3(\ddot{x})}{\Upsilon_3(x)}.$$

These coefficients contain adequate information about the frequencies ω_1 and ω_2, which are then computed as

$$\omega_{1,2} = \sqrt{\frac{c_1 \pm \sqrt{c_1^2 - 4c_2}}{2}}.$$

These frequency estimates are then used in conjunction with 2nd-order energy operators to develop estimates for the amplitude as follows:

$$a_{1,2}^2 = \frac{\omega_{2,1}^4(\Psi(x^{(3)}) - \omega_1^2\omega_2^2\Psi(\dot{x}))}{\omega_1^4\omega_2^4(\omega_1^2 - \omega_2^2)^2} - \frac{\omega_{2,1}^6(\Psi(\ddot{x}) - \omega_1^2\omega_2^2\Psi(x))}{\omega_1^4\omega_2^4(\omega_1^2 - \omega_2^2)^2}.$$

The GDE coefficients are time invariant quantities for sinusoidal signals and become slowly time-varying (lowpass) quantities for AM–FM signals.

3. DISCRETE–TIME ALGORITHM

Discrete–time two–component AM–FM signals are modeled instantaneously as two–component sinusoidal signals of the form

$$x_n = a_1 \cos(\Omega_1 n + \theta_1) + a_2 \cos(\Omega_2 n + \theta_2).$$

This mixture satisfies the following fourth-order generating difference equation

$$c_1(x_{n-1} + x_{n-3}) + c_2 x_{n-2} = -(x_n + x_{n-4}), \quad \text{where}$$

$$c_1 = -2(\cos\Omega_1 + \cos\Omega_2) \quad , \quad c_2 = 4\cos\Omega_1 \cos\Omega_2 + 2.$$

Evaluating the GDE at time instants n and $n+1$ and solving the 2×2 linear system of equations, we obtain

$$c_1 = \frac{\Upsilon_3(x_{n-3}) - \Upsilon_3(x_{n-1})}{\Psi(x_{n-1}) - \Psi(x_{n-2})}$$

$$c_2 = \frac{\Psi(x_n) - \Psi(x_{n-3})}{\Psi(x_{n-1}) - \Psi(x_{n-2})} + \frac{\Upsilon_4(x_{n-2}) - \Upsilon_4(x_{n-3})}{\Psi(x_{n-1}) - \Psi(x_{n-2})},$$

where Υ_k is the k^{th}-order discrete-time energy operator which we define as

$$\Upsilon_k(x_n) \quad \stackrel{\triangle}{=} \quad x_n x_{n+k-2} - x_{n-1} x_{n+k-1},$$

and Ψ is the discrete Teager-Kaiser energy operator [3]:

$$\Psi(x_n) \quad = \quad \Upsilon_2(x_n) \quad = \quad (x_n)^2 - x_{n+1} x_{n-1}.$$

The discrete-time EDM algorithm is then given by

$$\Omega_{1,2} = \cos^{-1}\left(\frac{-c_1}{4} \pm \frac{\sqrt{c_1^2 - 4c_2 + 8}}{4}\right)$$

$$a_{1,2}^2 \quad = \quad \frac{S_{2,1}^4(\Psi({\Delta_s}^3 x) - S_1^2 S_2^2 \Psi(\Delta_s x))}{S_1^4 S_2^4 (S_2^2 - S_1^2)^2} - \frac{S_{2,1}^6(\Psi({\Delta_s}^2 x) - S_1^2 S_2^2 \Psi(x))}{S_1^4 S_2^4 (S_2^2 - S_1^2)^2},$$

where $S_{1,2} = \sin(\Omega_{1,2})$, and

$$\Delta_s x = \frac{x_{n+1} - x_{n-1}}{2} \quad , \quad {\Delta_s}^m x = \Delta_s({\Delta_s}^{m-1} x).$$

Similar results for frequency estimation have been obtained in [10] for two–component sinusoidal signals in the context of AR signal modeling. The novelty of our procedure is its application to the demodulation of two–component AM–FM signals, the use of physically meaningful higher–order energy operators and amplitude estimation via energy equations instead of least squares.

If the signal of interest is corrupted by additive white Gaussian noise (AWGN), then the EDM algorithm is used in conjunction with the multiband filtering scheme proposed in [4] (multiband–EDM). The signal is filtered through a parallel bank of Gabor or FIR filters that sample the frequency domain densely. The center frequencies of the filters are uniformly spaced in frequency and the bandwidths are optimized for minimum harmonic distortion. An energy detector chooses the most active channel (largest short time mean Teager–Kaiser energy), while noise suppression is achieved by discarding all other less active channels.

4. PERFORMANCE OF THE EDM

Consider a mixture signal of two sinusoidally modulated and spectrally close FM signals

$$x[n] = \sum_{i=1}^{2} a_i \cos\left(\int_0^n \Omega_i[m] dm\right) \quad = \quad x_1[n] + x_2[n]$$

$$\Omega_i[n] = \Omega_{ci} + \Omega_{mi} \cos(\Omega_{fi} n + \theta_i) \quad , \quad i \ = 1 \ , \ 2.$$

The signal environment of the example is described in Table 3. Significant amount of spectral overlap is present as shown in Fig. 3(b). Fig. 3(d - h) describe the estimates of the multiband–ESA, the EDM, the CC–DPLL, the LMS and the ITD, while the demodulation errors (amp. and freq.) of the algorithms are tabulated in Fig. 3(c). The matrix system of the ITD algorithm develops singularities shown by the regions of constant estimates (Fig. 3(h)). Post smoothing of the estimates involves moving average filtering to remove beating at the carrier difference frequency and 9-pt median filtering to remove spikes. The EDM with post–smoothing is denoted as SEDM (smoothed EDM) and likewise for other methods. Carrier frequency and mean amplitude estimates are the mean of the instantaneous frequency and amplitude estimates over the duration of the signal.

Next we define some performance related parameters of the mixture signal. The *normalized carrier separation* (SEP) and *carrier to information bandwidth ratio* (CR/IB) parameters are

$$\text{SEP} = \frac{|\Omega_{c2} - \Omega_{c1}|}{2(\Omega_{fi} + \Omega_{ai} + \Omega_{mi})} \quad , \quad \text{CR/IB} = \frac{\Omega_{ci}}{\max(\Omega_{fi}, \Omega_{ai})}.$$

The SEP parameter measures the spectral separation between the components. The denominator of this parameter is the Carson bandwidth of the AM–FM signal, which under-estimates the actual bandwidth of the signal. The *signal to interference ratio* (SIR) and *carrier to frequency deviation ratio* (CR/FD) of the mixture are

$$\text{SIR} = 20\log\left(\frac{\sigma_{x1}}{\sigma_{x2}}\right) \quad \text{and} \quad \text{CR/FD} = \frac{\Omega_{ci}}{\Omega_{mi}},$$

where σ_{xi} are the standard deviations of the components of the mixture. The SIR parameter measures the power of the first component relative to the second, while the CR/FD parameter measures the strength of signal modulations.

As the CR/FD parameter increases, the strength of signal modulation relative to the interference decreases and the demodulation error increases as depicted in Fig. 1(a) for three different cases. As the CR/IB parameter increases, the FM mixture approaches a stationary sinusoidal mixture, resulting in a decrease in the demodulation error (Fig. 1(b)). As the SIR increases, the power of the first component relative to the second component increases. The frequency estimation section of the EDM is independent of the SIR parameter (Fig. 1(c)). For large SIR parameters, the stronger component dominates the mixture, resulting in a decrease in the amplitude demodulation error of the stronger component, while that of the weaker one rapidly increases (Fig. 1(d)). When the SEP parameter decreases the spectral overlap between the components increases and

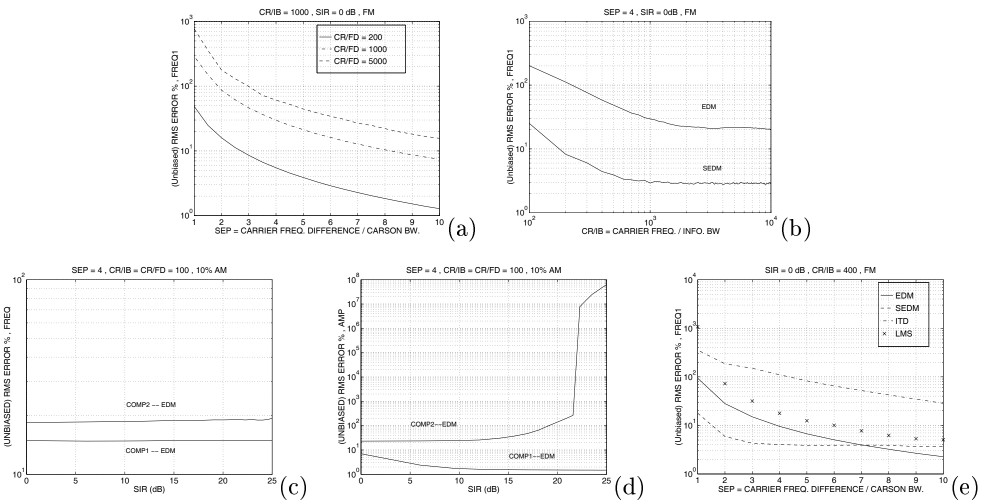


Figure 1. (a) Effect of CR/FD**, (b) effect of** CR/IB **on frequency demodulation, (c,d) effect of** SIR **on frequency and amplitude demodulation and (e) effect of** SEP **on frequency demodulation (by averaging over FM modulation indices** $\in \{0.1 - 1\}\%$**)**

beyond a limit the energy equations of the EDM become ill-conditioned and finally singular as described in the case of the unmodulated sine mixture by

$$\Upsilon_3(x) = 0 \quad , \quad \Psi(x_{n-1}) - \Psi(x_{n-2}) = 0,$$

causing an increase in the demodulation error (Fig. 1(e)).

5. APPLICATION TO SPEAKER SEPARATION

In this section, the EDM algorithm was applied on two–component voice–modulated FM signals. Two–component FM signals, with specified parameters are created using two voice signals. The created two–component mixture is then demodulated using the EDM to obtain estimates of the message signals. This experiment was performed for equal amplitudes, 6% FM modulation and for carrier separations of $\{400, 200, 100, 50, 33, 25, 17\}\%$. The RF bandwidth of each component is approximately 12 kHz and the sampling frequency is 400 kHz. The first carrier is maintained at 120 kHz, while the other is variable.

The estimates of the message signals are clearly intelligible for separations of $\{400, 200, 100\}\%$ and the estimation errors are small in magnitude compared to the messages. The energy in the spectral overlap or the interaction between the components at these separations is negligible. For smaller separations of $\{55, 33, 25\}\%$, the interference becomes audible and the estimation error signals are larger, while the message signals are still intelligible.

As the carrier separation becomes smaller, the magnitude of the estimation errors becomes comparable to that of the messages. The energy in the spectral overlap between the components becomes large and the EDM starts to encounter singularity problems which manifest themselves as beating in the estimates. For further decrease in separation, the beating becomes stronger than the message signal and eventually swamps out the messages.

REFERENCES

[1] P. Maragos, J. F. Kaiser, and T. F. Quatieri, "On Separating Amplitude from Frequency Modulations Using Energy Operators," in *Proc. ICASSP92*, vol. 2, pp. 1 – 4.

[2] P. Maragos, J. F. Kaiser, and T. F. Quatieri, "Energy Separations in Signal Modulations with Application to Speech Analysis," *IEEE Trans. on Sig. Proc.*, vol. 41, pp. 3024 – 3051, Oct 1993.

[3] J. F. Kaiser, "On a simple algorithm to calculate the energy of a signal," in *Proc. ICASSP90*, pp. 381 – 384.

[4] A. C. Bovik, P. Maragos, and T. F. Quatieri, "AM–FM Energy Detection and Separation in Noise Using Multiband Energy Operators," *IEEE Trans. on Sig. Proc.*, vol. 41, pp. 3245 – 3265, Dec 1993.

[5] L. J. Griffiths, "Rapid Measurement of Digital Instantaneous Frequency," *IEEE Trans. on ASSP*, vol. 23, pp. 207 – 222, April 1975.

[6] R. Kumaresan, A. G. Sadasiv, C. S. Ramalingam, and J. F. Kaiser, "Instantaneous Nonlinear Operators For Tracking Multicomponent Signal Parameters," in *Proc. of the* 6^{th} *SP Workshop on Statistical Sig. and Array Proc.*, pp. 404 – 407, Oct 1992.

[7] C. L. Dimonte and K. S. Arun, "Tracking the Frequencies of Superimposed Time–Varying Harmonics," in *Proc. ICASSP90*, vol. 2, pp. 2539 – 2542.

[8] J. N. Bradley and R. L. Kirlin, "Phase–Locked Loop Cancellation of Interfering Tones," *IEEE Trans. on Sig. Proc.*, vol. 41, pp. 391 – 394, Jan 1993.

[9] P. Maragos and A. Potamianos, "Higher Order Differential Energy Operators," *IEEE Sig. Proc. Letters*, vol. 2, pp. 152 – 154, Aug 1995.

[10] T.-H. Li and B. Kedem, "Iterative Filtering for Multiple Frequency Estimation," *IEEE Trans. on Sig. Proc.*, vol. 42, pp. 1120 – 1132, May 1994.

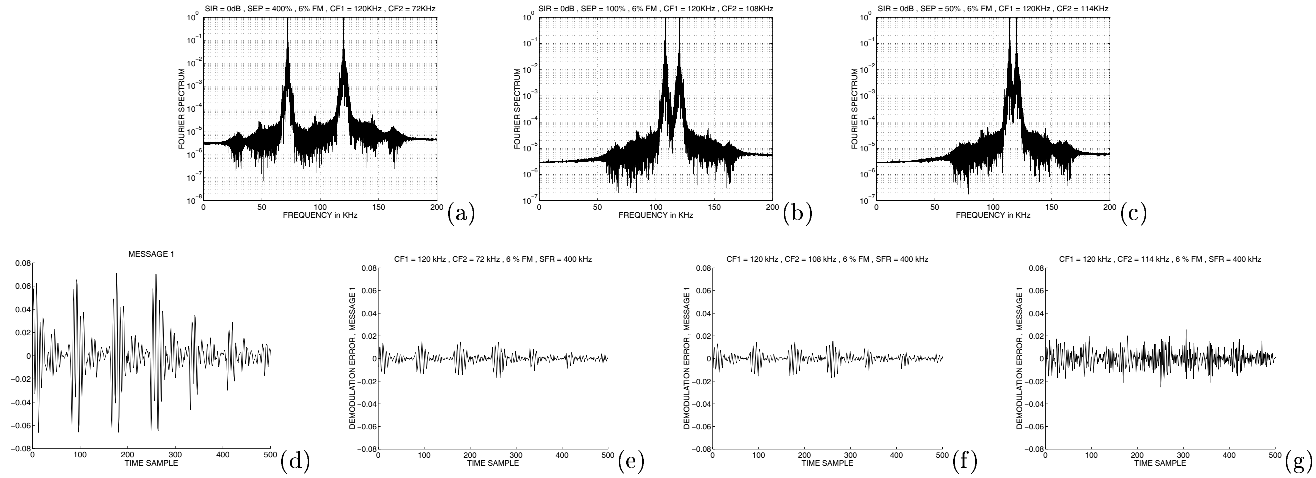
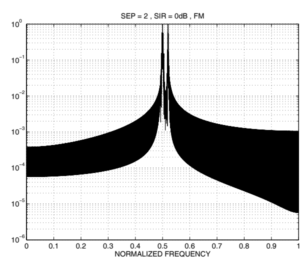


Figure 2. (a,b,c) are the spectra of the composite signal for normalized carrier separations of $\{400, 100, 50\}\%$. (d,e,f,g) are the original voice signal and the demodulation errors for the first signal component at normalized carrier separations of $\{400, 100, 50\}\%$, pertaining to the EDM–related voice experiments.

Parameter	Value
Ω_{c1}	$\frac{\pi}{2}$
θ_i	$0, \frac{\pi}{2}$
SEP	2
SIR	0 dB
κ_i	0 , 0
CR/FD	200 , 400
CR/IB	200 , 400

(a)



(b)

Algorithm	Freq. 1	Freq2.	Amp. 1	Amp. 2
SMESA	14	6.4	3.1	1.94
EDM	24.9	52.9	3.4	7.4
SEDM	11.6	6.0	0.8	0.9
SLMS	29.6	37.8		
SITD	Singular	Singular	Singular	Singular
SCC-DPLL	17.8	19.6	5.3	5.5

(c)

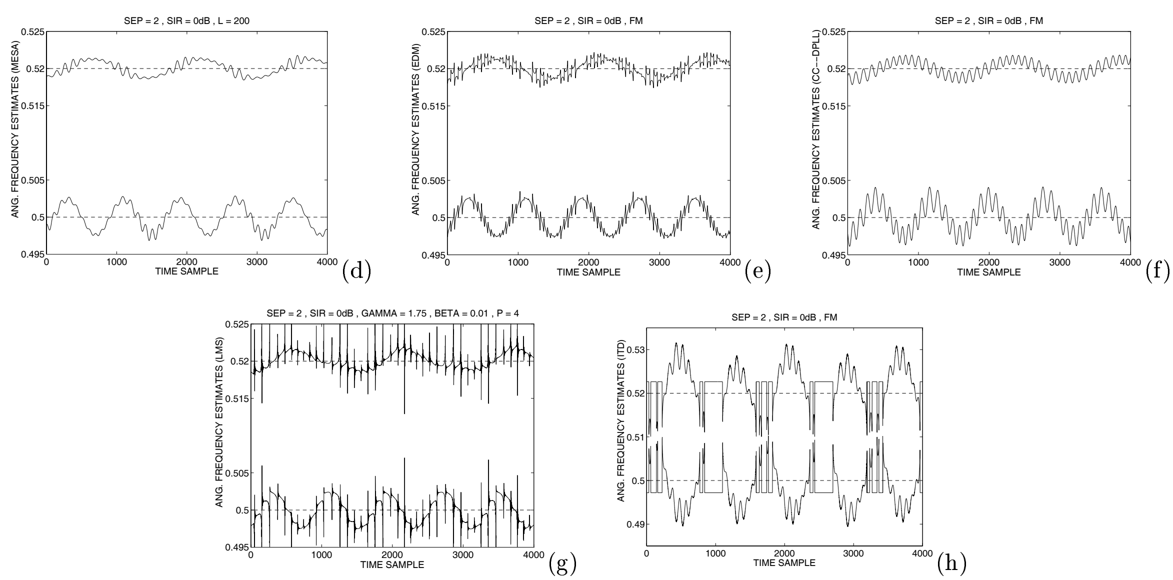


Figure 3. (a) Signal environment of example, (b) Fourier spectral magnitude of the composite signal, (c) percentage RMS frequency and amplitude demodulation errors of the algorithms, angular frequency estimates of (d) the multiband–ESA (using FIR filters of length 200), (e) the EDM algorithm, (f) the CC–DPLL algorithm, (g) the LMS algorithm and (h) the ITD algorithm (constant regions are regions of singularity) as fractions of π.