Robust FM demodulation of discrete-time signals using least squares differential ratio

J. Epps, E. Ambikairajah and T. Thiruvaran

A technique for the robust demodulation of discrete FM signals in the presence of additive noise is presented. Based upon least squares, the proposed technique is shown to improve the accuracy of instantaneous frequency estimation by 17 to 79% compared with the popular discrete energy separation (DESA), smooth DESA and Hilbert transform-based algorithms, e.g. noisy signals at low SNRs.

Introduction: Estimation of the instantaneous frequency has recently regained attention in the speech processing community as researchers look for alternatives to spectral magnitude-based models of the speech signal. Among possible approaches to FM demodulation of discrete signals with time-varying amplitude, one is to derive an expression for the instantaneous frequency of the signal that eliminates the amplitude. The advantage of this approach is that inaccuracies in amplitude estimates then do not degrade the accuracy with which the instantaneous frequency can be estimated, as can occur in methods such as the Hilbert transform [1]. The discrete energy separation algorithm (DESA) [2] is one example of an expression that eliminates the amplitude, which has been applied successfully to speech recognition recently [3] and shown to outperform other FM demodulation schemes for some applications [4, 5]. Unfortunately the performance of DESA in FM demodulation degrades quite severely in the presence of additive noise, detracting from its potential in such applications. In this Letter, we examine how least squares methods can be applied to an FM demodulation approach of this type to improve demodulation performance in noise.

Derivation of differential ratio demodulation method: Among the simplest means for eliminating the amplitude of a periodic signal for the purposes of frequency estimation is that based on the observation that for a continuous-time sinusoidal signal

$$x(t) = A\sin(\omega t + \varepsilon) \tag{1}$$

with amplitude A and analogue frequency ω, both assumed constant,

$$\ddot{x}(t) = -\omega^2 x(t) \tag{2}$$

Thus ω can be estimated directly via the ratio of $\ddot{x}(t)$ to $x(t)$ By contrast, the DESA method eliminates the amplitude A through the relationship [2]

$$\begin{aligned}\Psi(x(t)) &= \dot{x}^2(t) - x(t)\ddot{x}(t)\\ &= A^2\omega^2\end{aligned} \tag{3}$$

from which ω can be estimated directly via the ratio of $\Psi(\dot{x}(t))$ to $\Psi(x(t))$. The relationship in (2) is of course familiar from physics as part of the solution to the wave equation, and has been identified for use in instantaneous frequency estimation for some time [6]. If we now examine this relationship for a discrete signal, $x[n]$ by analogy with (2) we have

$$d_2[n] = -\hat{\theta}^2 x[n] \tag{4}$$

where $\hat{\theta}$ is the digital frequency estimate and $d_2[n]$, the second derivate of $x[n]$, is approximated here using a first-order central difference

$$d_2[n] = x[n] - 2x[n-1] + x[n-2] \tag{5}$$

The digital frequency can thus be estimated using

$$\hat{\theta}^2 x[n] = -d_2[n] = -x[n] + 2x[n-1] - x[n-2] \tag{6}$$

We refer to this proposed frequency estimation technique as the differential ratio method. The relationship between $\hat{\theta}$ and the true digital frequency θ can be found by taking the z-transform of (6) and substituting $z = e^{j\theta}$,

$$\begin{aligned}\hat{\theta}^2 &= -(1 - z^{-1})^2\\ |\hat{\theta}|^2 &= 4\sin^2\left(\frac{\theta}{2}\right)\end{aligned} \tag{7}$$

Hence, as a result of the approximation to differentiation, a warping of $\theta = 2\arcsin(\hat{\theta}/2)$ needs to be applied to the estimate $\hat{\theta}$ to obtain the true frequency θ for a clean signal $x[n]$. Computationally, the arcsin transformation is not very attractive. Taylor series represent an alternative, however informal empirical work indicates that a very high order is needed for frequencies close to $\hat{\theta} = \pi$. For low frequencies where $\hat{\theta} \simeq \theta$, the arcsin warping can be neglected altogether. A major drawback of the differential ratio method of frequency estimation is the division by $x[n]$ implied by (6), which besides creating problems when $x[n]$ is zero, produces large errors in frequency estimation in the presence of noise.

Robustness enhancement of differential ratio demodulation method: Here, we address this sensitivity to noise using a least squares approach. Rewriting (5) in matrix-vector notation over a window of length N,

$$\mathbf{d}_2 \simeq -a\mathbf{x} \tag{8}$$

where

$$\mathbf{d}_2 = \begin{bmatrix} x[n] - 2x[n-1] + x[n-2] \\ \vdots \\ x[n-N+1] - 2x[n-N] + x[n-N-1] \end{bmatrix} \tag{9}$$

$$\mathbf{x} = \begin{bmatrix} x[n-1] \\ \vdots \\ x[n-N] \end{bmatrix} \tag{10}$$

and $a = \hat{\theta}^2$. Then

$$|a| = -[\mathbf{x}^T\mathbf{x}]^{-1}\mathbf{x}^T\mathbf{d}_2 \tag{11}$$

where T denotes vector transpose. Note that $\boldsymbol{x}^T\boldsymbol{x}$ is a scalar, so no matrix inversion is required. Thus a more robust frequency estimate than the differential ratio technique can be obtained as

$$\hat{\theta}_{LS} = 2\arcsin\left(\frac{\sqrt{q}}{2}\right) \tag{12}$$

Note that in practice a is nearly always positive (always for SNR > 20 dB and 99.5% of the time for SNR = 0 dB in our experiments), however $|a|$ can be used in place of a where required. In principle, a similar least squares approach to that of (8)–(11) can be employed with the DESA method, however informal empirical results for this approach were not promising.

Evaluation: Since the main purpose of deriving a robust instantaneous frequency estimation technique is for FM demodulation, in this Section we examine the performance of several estimation techniques for the reference AM-FM signal used in [2],

$$s[n] = \left(1 + 0.5\cos\left(\frac{\pi n}{50}\right)\right)\cos\left(n\theta_c + 4\sin\left(n\theta_m + \frac{\pi}{4}\right)\right) \tag{13}$$

where θ_c and θ_m are the carrier and message frequencies, respectively.

First, the least squares window length N of the least squares differential ratio technique was varied for a range of different message frequencies, in order to determine the length required for accurate instantaneous frequency estimation, in terms of mean-square error. The second experiment aimed to compare the differential ratio and least squares differential ratio with existing approaches for noisy signals. Here the DESA [2], smooth DESA [4] and Hilbert transform [1] algorithms were also implemented, and all five algorithms were used to demodulate a noisy signal $x[n] = s[n] + w[n]$, where $w[n]$ is a normally distributed white noise process and the parameter settings $\theta_c = \pi/4$, $\theta_m = \pi/50$, and $N = 5$ were used. The Hilbert transform method was implemented with an FIR filter of length 19, similarly to the method referred to as 'short HTSA' in [4]. The frequency estimates from all algorithms were lowpass filtered using a rectangular window of length 100, for similar reasons to the smoothing approaches applied in [3]. These frequency estimation measurements were repeated over different SNRs, and the mean-square error from the true instantaneous frequency was evaluated for each case, as shown in Fig. 1. For each approach, mean and standard deviation normalisation was applied to the frequency estimates and any phase difference from the instantaneous frequency of the true signal in (13) was removed, to permit fair comparison between them.