A JOINT VITERBI ALGORITHM TO SEPARATE COCHANNEL FM SIGNALS

Jon Hamkins

Jet Propulsion Laboratory
Mail Stop 238-420
4800 Oak Grove Drive
Pasadena, CA 91109-8099
email: hamkins@jpl.nasa.gov

ABSTRACT

This paper presents a method for separating cochannel FM signals. We show that the Viterbi algorithm, traditionally limited to estimation of digital sequences, can jointly track analog FM signals by separately quantizing the derivatives of their instantaneous frequencies. We employ per-survivor processing in the trellis to estimate unknown channel effects. The approach works well when the signal to interference ratio (SIR) is less than or equal to zero, in contrast to conventional interference suppression algorithms that degrade as SIR approaches zero and fail catastrophically when SIR < 0. Comparisons of mean squared error (MSE) between the estimates and the true signals are given for varying SIR, SNR, Doppler offsets, and frequency deviations. The same approach can also be used for any other continuous phase modulation scheme, such as continuous-phase frequency-shift keying (CPFSK).

1. INTRODUCTION

A single phase-locked loop (PLL) or phase discriminator can effectively demodulate an FM signal because the signal has a constant envelope and an instantaneous frequency that is proportional to the message signal. However, these conventional techniques can suffer severe degradation when a second cochannel FM signal is added to the receiver input, because the envelope is no longer constant and the instantaneous frequency is not proportional to either of the modulating signals or their sum. The PLL output in this case contains large, inband spikes and the output is unintelligible [7]. As a result, a number of different receivers have been developed to combat cochannel interference [2, 4, 5, 7, 9]. Such work has made important contributions, but each is missing one or more of the following desirable attributes:

1. The estimation of both dominant and subdominant signal components.
2. The estimation of unknown channel parameters with Per-Survivor Processing (PSP) in a trellis algorithm, instead of with weaker decision directed estimators.
3. The applicability to both digital (CPFSK) and analog (FM) continuous phase modulation schemes.

In this paper, we present a receiver that contains all three features; consequently, it exhibits improved demodulated voice quality for cochannel FM signals. The receiver uses a method of Cahn [1] (and later, [6]) to transform the phase tracking or FM demodulation problem into a discrete sequence estimation problem that can be solved with the usual Viterbi algorithm. Unlike Cahn's work, however, we use a joint trellis capable of tracking both dominant and subdominant parts of a cochannel signal, and we estimate unknown channel effects with PSP.

2. SIGNAL MODEL

For simplicity, we assume that there are exactly two cochannel signals. Extension to additional signals is straightforward. The complex baseband representation of the sampled received signal is

$$r[k] = A_1[k]e^{j\theta_1[k]} + A_2[k]e^{j\theta_2[k]} + N[k], \qquad (1)$$

where $A_i[k]$ and $\theta_i[k]$ is the amplitude and phase of the ith signal at time kT_s, respectively, where T_s is the sampling period, and where $N[k]$ is a complex noise process. The amplitude is assumed to vary much slower than the phase, which is further decomposed as

$$\theta_i[k] = \omega_i kT_s + \phi_i + k_i \int_0^{kT_s} m_i(s)\, ds, \qquad (2)$$

where for the ith signal, ω_i is an offset carrier frequency in radians/second, ϕ_i is an initial phase offset in radians,

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

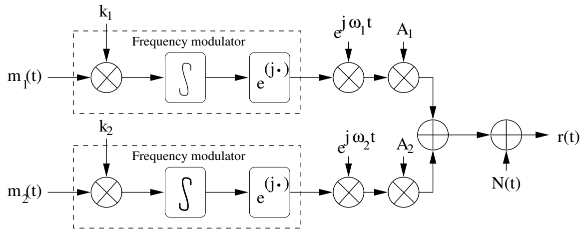


Figure 1: The signal model for two cochannel signals.

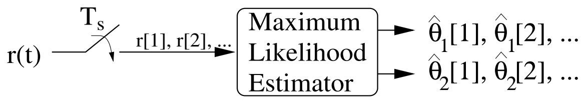


Figure 2: A maximum likelihood joint phase estimator.

k_i is the frequency deviation in radians/second, and $m_i(s)$ is the message waveform. This is illustrated in Figure 1. This paper assumes that $-1 \leq m_i(s) \leq 1$ for all $s > 0$ and $i \in \{1, 2\}$. Note that through suitable restrictions on $m_i(s)$ involving symbol times and levels, digital modulating schemes such as CPFSK are implicitly allowed in Equation (2).

3. JOINT VITERBI ESTIMATION WITH PSP

Given the vector $Y = (r(1), r(2), \ldots, r(N))$, the maximum likelihood estimate of $(\theta_1[k], \theta_2[k])$ is

$$(\hat{\theta}_1[k], \hat{\theta}_2[k])_{ML} = \arg\left\{ \max_{\theta_1[k], \theta_2[k]} p_{\theta_1[k], \theta_2[k]}(Y) \right\},$$

where $p_{\theta_1[k],\theta_2[k]}(Y)$ is the joint probability density function of $(\theta_1[k], \theta_2[k])$ evaluated at Y. This is illustrated as a black box in Figure 2. Since $A_1[\cdot]$, $A_2[\cdot]$, $m_1[\cdot]$, and $m_2[\cdot]$ are in general each unknown, computing the maximum likelihood estimate of $(\theta_1[k], \theta_2[k])$ involves a joint maximization over at least four variables. In practice, the joint probability density function is not known exactly, and in any case such a maximization would be too complex to implement.

To make the problem tractible, we start by modeling the uncountably infinite possibilities for the trajectory of $\theta_i[k]$ by a countable number. This can done by quantizing $\theta_i[k]$ to one of a finite number of phases; however, better performance can be had by instead quantizing the second derivative of the phase at time kT_s, denoted by $\theta_i''[k]$. This was the approach taken by Cahn for the case of tracking the phase of a single signal [1]. The true second derivative, determined from the continuous version of Equation (2), is given by $\theta_i''[k] = k_i m_i'[k]$, where $m_i'[k]$ is the derivative of the message waveform evaluated at time kT_s, and can be any real number. The estimator, however, assumes that $k_i m_i'[k]$ can take on only the values C or $-C$, where C is a constant. Of course, this is not the case, but when integrated twice and for a suitable sample rate, little fidelity is lost in the approximation.

Each state in the the kth column of the trellis stores the estimates $\hat{\theta}_i''[k] \in \{-C, C\}$, $\hat{\theta}_i'[k]$, $\hat{\theta}_i[k]$, and $\hat{A}_i[k]$ for $i \in \{1, 2\}$, as well as the accumulated metric for that state. The binary representation of the index of the state indicates the estimate $\hat{\theta}_i''[k]$. For example, the 46th state in the kth column of a 64-state trellis has binary representation 101110, and indicates that

$$(\hat{\theta}_1''[k], \hat{\theta}_1''[k-1], \hat{\theta}_1''[k-2], \hat{\theta}_2''[k], \hat{\theta}_2''[k-1], \hat{\theta}_2''[k-2]) = (C, -C, C, C, C, -C).$$

We can efficiently compute the four outgoing branches of each state based on the index of the state. In the example, state (101110) at time k can go to any state of the form (01X10Y) at time $k+1$, where $X, Y \in \{0, 1\}$.

Once $\theta_i''[k]$ is estimated as $\hat{\theta}_i''[k] = \pm C$, each state uses a truncated Taylor series to estimate the instantaneous frequency and the phase:

$$\begin{aligned} \hat{\theta}_i'[k] &= \hat{\theta}_i'[k-1] + \hat{\theta}_i''[k-1]T_s \\ \hat{\theta}_i[k] &= \hat{\theta}_i[k-1] + \hat{\theta}_i'[k-1]T_s + \hat{\theta}_i''[k-1]T_s^2/2. \end{aligned}$$

The amplitude estimates are computed with a gradient descent algorithm. A typical gradient technique would update a single estimate of $(A_1[k], A_2[k])$ by computing the derivative of the accumulated metric of a tentative winning state with respect to the estimate $(\hat{A}_1[k], \hat{A}_2[k])$. The trellis would then use this estimate in every state for the computation of the branch metric. We abandon this approach and instead keep a separate amplitude estimate at *each* state of the trellis. This is called per-survivor processing (PSP) [8], and offers improved performance because, unlike the single amplitude estimator approach, when a particular path through the trellis is chosen, the amplitude estimates used in that path are optimized for that path. In other words, there is no penalty when a tentative path does not turn out to be the path ultimately chosen.

Thus, the sequence of signs for $\hat{\theta}_i''[\cdot]$ determines the sequence $\hat{\theta}_i[k]$, which together with amplitude estimates gives rise to an estimated remodulated signal of

$$\hat{r}[k] = \hat{A}_1[k]e^{j\hat{\theta}_1[k]} + \hat{A}_2[k]e^{j\hat{\theta}_2[k]}. \tag{3}$$

Since the noise is assumed to be AWGN, the maximum likelihood estimation of $(\theta_i''[k], \theta_i''[k])$ is that which minimizes the Euclidean distance $\sum_{i=1}^{k} ||r[l] - \hat{r}[l]||^2$. A joint Viterbi algorithm is used to trace through the trellis and obtain the optimum sign sequence for each of $\theta_1''[\cdot]$ and $\theta_2''[\cdot]$.

There are a number of design issues for which space does not permit a full description. See [3] for details. The choice of the constant C affects performance and must be carefully chosen. More than the two levels C and $-C$ may be used to increase performance; for example, an additional level of zero could be added. Also, the appropriate size of the trellis must be determined. The size of the trellis here is determined by the "memory" of $\theta''_i[k]$, just as the trellis size of a maximum likelihood sequence estimator in an ISI environment is determined by the memory of the channel. We have found that devoting three "bits" (sign assignments) of memory to each of $\theta''_1[k]$ and $\theta''_2[k]$ works well. This results in a $2^{2\cdot 3} = 64$ state trellis. If it is known that digital modulation schemes are being used, this can be easily modeled in the trellis and may result in improvement over the analog quantization presented here. Finally, internal interpolation of the received signal to obtain a higher sampling rate can also improve performance, as can appropriate pre- and post-processing filters, which reduce the effects of out-of-band noise.

4. SIMULATED PERFORMANCE

The performance results are stated in terms of the mean squared error (MSE) between the true sampled message signal $m_i[k]$ and the estimate $\hat{m}_i[k]$, normalized by the signal power:

$$\text{Normalized MSE} = \frac{\sum_{l=1}^{M} \left(\hat{m}_i[l] - m_i[i]\right)^2}{\sum_{l=1}^{M} m_i[l]^2}, \qquad (4)$$

This metric is somewhat problematic because it fails to capture some important information. For example, an algorithm which contains a rare spike but otherwise perfectly tracks the phase may have a higher relative error than a generally noisy phase tracker with no large spikes. A better test may be a qualitative assessment of the output audio. As a rule of thumb, we view an algorithm as "working" if the normalized MSE is less than about 1.0. This may seem to be a very liberal rule in view of the fact that the all zero output achieves this MSE. However, we have found that in nearly all cases, when the algorithms achieve a relative error of about 1.0, or even a little higher, they lock on to significant portions of the intended signal and have good voice quality.

The normalized MSE performance for a cochannel signal is shown in Table 1 for varying SIR, SNR, frequency deviations and Doppler offsets. It compares the joint Viterbi algorithm (using PSP estimates), a PLL, and a differential phase detector (DPD) or discriminator. In all cases reported, the sampling rate is $T_s = 1/65536$ and the frequency deviation of the first signal is $k_1 = 24000\pi$ (i.e., 12 kHz). The two signals consist of simulated voice waveforms, each with a 3.7kHz bandwith and 3 seconds in length. This allows for the processing of about 200,000 samples in each case. The signal to interference ratio, defined by SIR $= 20\log_{10}(A_1/A_2)$, was varied between 2dB and 6dB. The SNR, defined as the ratio of the power of the first signal to

Table 1: Comparison of cochannel FM signal estimators.

SIR	SNR	k_2	ω_2	Normalized MSE		
dB	dB	kHz	kHz	Viterbi*	PLL†	DPD†
6	∞	12	0	0.12/0.68	0.07	0.41
6	∞	12	1	0.15/0.65	0.08	0.41
6	∞	24	0	0.13/0.51	0.09	0.79
6	∞	24	1	0.15/0.47	0.09	0.82
6	10	12	0	0.55/2.09	0.29	2.46
6	10	12	1	0.37/1.68	0.30	2.51
6	10	24	0	0.31/1.37	0.50	2.89
6	10	24	1	0.41/1.05	0.47	2.90
2	∞	12	0	0.42/0.65	0.18	1.58
2	∞	12	1	0.46/0.65	0.19	1.65
2	∞	24	0	0.61/0.34	0.20	2.79
2	∞	24	1	0.56/0.40	0.20	2.85
2	10	12	0	0.86/1.01	–	5.44
2	10	12	1	0.98/1.03	–	5.77
2	10	24	0	1.08/0.54	0.74	6.88
2	10	24	1	1.21/0.54	–	6.84

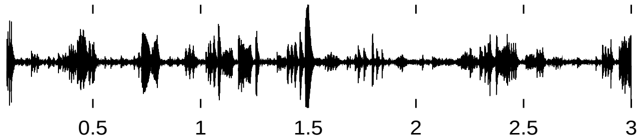
* Performance for dominant and subdominant signals
† Performance for dominant signal only

the noise power, was varied between 10dB and ∞. The subdominant frequency deviation was varied between 12kHz and 24kHz. A Doppler offset between the signals was varied between 0 and 1kHz.

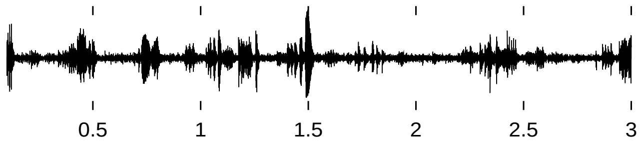
This is among the first work which demonstrates that a subdominant FM signal may be captured when a cochannel FM signal at twice the power is interfering. Thus, by switching one's perspective of the "first" and "second" signals, the SIR's of 2dB and 6dB may actually be considered as SIR's of -2dB and -6dB, and still with excellent results for a variety of transmission parameters. Conventional techniques do not attempt to estimate the subdominant signal and can only hope to *reject* the interference, at best. Thus, such techniques have no hope for operating in situations in which SIR ≤ 0. Indeed, for the PLL and DPD, Table 1 can only report the MSE error with respect to their lone output: their estimate of the dominant signal.

An example of the estimated voice waveforms is shown in Figure 3. The waveforms were taken from the simulation given in the third row of Table 1. As can be seen, both the PLL and the joint Viterbi algorithm effectively capture the dominant signal, and the DPD is somewhat affected by the cochannel interference. The subdominant signal is accurately tracked by the joint Viterbi algorithm.

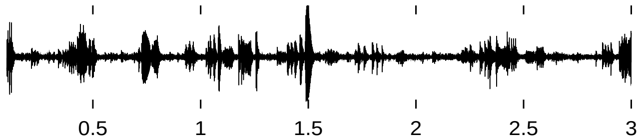
Preliminary testing of cochannel FM signals recorded from the airwaves has been performed. The recorded signals had smaller frequency deviations, a much harder test, and consequently we were unable to verify the excellent



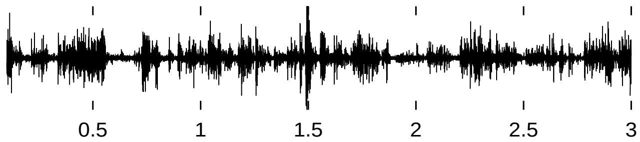
(a) Dominant (original) speech signal.



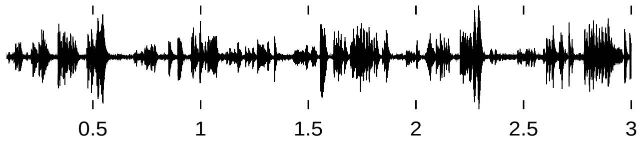
(b) Joint Viterbi estimate of dominant signal.



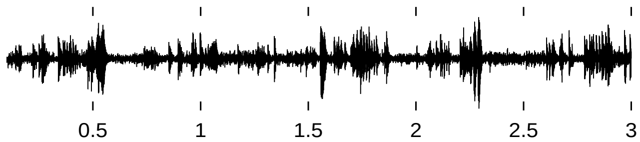
(c) PLL estimate of dominant signal.



(d) DPD estimate of dominant signal.



(e) Subdominant (original) speech signal.



(f) Joint Viterbi estimate of subdominant signal.

Figure 3: Estimated dominant and subdominant message waveforms .

subdominant performance obtained above. However, the performance of the real signals agrees with simulated performance done at the lower frequency deviations.

5. CONCLUSIONS

The joint Viterbi phase tracker separates cochannel signals for a variety of SIR, SNR, Doppler offsets, and frequency deviations. It works for FM, CW, and CPFSK signals. Its estimate of the dominant signal is usually slightly worse than a conventional PLL, but when the desired signal is not dominant it can still properly track the phase, whereas the PLL can track only the dominant signal. When the constituent cochannel signals have equivalent power, the PLL and DPD break down and the advantages of the joint Viterbi approach become more apparent.

This joint estimator may be extended to more than two interfering signals with the addition of more trellis states. Since the trellis size grows exponentially in the number of interfering signals, there is a practical limit to the model, however.

Acknowledgements: The author expresses thanks to Ed Satorius and Norm Lay for for many helpful discussions.

6. REFERENCES

[1] Charles R. Cahn. Phase tracking and demodulation with delay. *IEEE Trans. Inform. Theory*, IT-20(1):50–58, January 1974.

[2] K. Giridhar, , John J. Shynk, Amit Mathur, Sujai Chari, and Richard P. Gooch. Nonlinear techniques for the joint estimation of cochannel signals. *IEEE Trans. Commun.*, 45(4):473–484, April 1997.

[3] Jon Hamkins. Joint phase tracking of co-channel signals using a Viterbi algorithm. Report of contracted work at JPL, 1997.

[4] Brad A. Hedstrom and R. Lynn Kirlin. Co-channel signal separation using coupled digital phase-locked loops. *IEEE Trans. Commun.*, 44(10):1373–1384, October 1996.

[5] Kazuhito Ishida, Ikuo Oka, and Ichiro Endo. Cochannel interference cancelling effects of MSE processing and Viterbi decoding in on-board processing satellites. *IEEE Trans. Commun.*, COM-34(10):1049–1053, October 1986.

[6] Odile Macchi and Louis L. Scharf. A dynamic programming algorithm for phase estimation and data decoding on random phase channels. *IEEE Trans. Inform. Theory*, IT-27(5):581–595, September 1981.

[7] Archibald M. Pettigrew. The amplitude-locked loop, the theory and operation of the FM 201/5 high performance FM demodulator. Ampsys company document, 1996.

[8] Riccardo Raheli, Andreas Polydoros, and Ching-Kae Tzou. Per-survivor processing: a general approach to MLSE in uncertain environments. *IEEE Trans. Commun.*, 43(2/3/4):354–364, February/March/April 1995.

[9] Michael J. Ready and Sujai Chari. Demodulation of cochannel FSK signals using joint maximum likelihood sequence estimation. In *The Twenty-Seventh Asilomar Conference on Signals, Systems & Computers*, volume 2, pages 1412–1416, Pacific Grove, California, November 1993.