A RATE 3/4 TCM DECODER FOR LINE SPECTRAL PAIRS USING MAP INFORMATION

W.N. Farrell and W.G. Cowley

Institute for Telecommunications Research
Department of Electronic Engineering
University of South Australia

ABSTRACT – This paper describes two schemes which combine Line Spectral Pair (LSP) source information with a Trellis–Coded Modulation (TCM) scheme. The proposed schemes use a rate 3/4 TCM scheme which is used to transmit LSP values across a channel. The decoder proposed is a Viterbi decoder which uses LSP Transition information to help correct errored paths and choose the most probable set of LSPs. Two variations of this decoder are discussed. By using knowledge of LSP transitions and expected noise variance, a–posteriori probabilities can be used as the metrics for branches in the trellis. This provides for more robust transmission over low SNR channels. Results show that there is significant gain in using either of the proposed methods.

INTRODUCTION

Line Spectral Pairs (LSPs) are used to describe the vocal tract filter for many different types of speech coders. LSPs are quite sensitive parameters and when errored produce annoying distortion such as beeps and "water droplet" sounds, which can often reduce intelligibility. Therefore it is important to provide some sort of error protection to the LSPs in order to minimise any distortion. Standard error protection codes can be employed, such as FEC, however it is much more effective to use some joint source and channel codec to exploit the characteristics of the LSP parameters and gain extra performance. Some work has been done in this area [2][3], each employing different methods of quantising and protecting LSP parameters.

Proposed is a rate 3/4 2xQPSK Trellis–Coded Modulation (TCM) decoder which uses the transition probabilities of the LSPs to help the Viterbi decoder select the correct path of the trellis. Optimum detection techniques have been presented in [4], which state that extra performance can be gained if residual redundancy of the source encoder is used at the decoder. For this work, the LSPs are quantised using the DoD Federal Standard CELP [1], which uses 34 bits for the 10 LSPs. This scalar quantisation has a large amount of residual redundancy present that can be exploited. Each codeword of the rate 3/4 TCM is mapped to the 3 most significant bits of each LSP. The four bits left over from the 4 4–bit LSPs are modulated using standard QPSK as they have lower sensitivity. There are two versions that are presented and tested. The first scheme uses the transition matrices for removing illegal branches in the trellis. If an LSP transition has a probability of zero, then the corresponding branch in the trellis is disallowed. This can help the Viterbi decoder choose the correct path if errors move an LSP to an illegal value. Zero probability transitions correspond to both non–monotonic LSP transitions and low value to high value LSP transitions. For this scheme, the transition probabilities are not included in the branch metric. The second, more optimum method, calculates the maximum a–posteriori (MAP) probability for each branch in the trellis. The a–posteriori probabilities are a combination of the Euclidean distance modulation information and the joint probability distribution of an LSP transition, which is contained in transition matrices. A requirement of the scheme is that an estimate of the channel noise variance is required to scale the branch metric.

The combination of a channel codec with a source coder is important. It is possible to use standard PSK modulation and still use a Viterbi decoder to help find the most probable set of LSPs sent. However, there is an advantage of using trellis–coded modulation in that the trellis code provides extra distance between possible LSP transitions. Also at the stage where the trellis starts to break down, the source information helps by weighting each branch and disallowing illegal branches. The result is that the trellis decoder performs much better for the very high BER channels. This is confirmed by the results presented.

RATE 3/4 TRELLIS CODED MODULATION (TCM)

Required is a TCM coder which provides good performance in poor channels and matches well to the LSP values. From [5] & [6], multiple symbol TCM can work well by using the extra dimensionality of the multiple symbols. The design used for this work is a simple, rate 3/4, 4–state trellis coder [7] which uses two QPSK symbols. One bit is used to describe the path through the trellis and details which modulation subset is to be used, whilst the other two bits index into the subset. The squared Euclidean distance between the uncoded bits is equal to four, while the trellis bit has a squared Euclidean distance of ten (due to the trellis). This means the trellis bit will perform very well asymptotically but worse than the other bits at low channel SNR due to wrong paths through the trellis. In figure 1, the modulation values correspond to the placements on the QPSK constellation.

The MSB of each LSP is mapped to the trellis bit of the TCM. The reason for this is two–fold; the trellis bit provides the best performance, except for only extremely bad channels and the MSB provides the most restriction due to LSP transitions. It is this restriction that helps the trellis choose a more probable path.

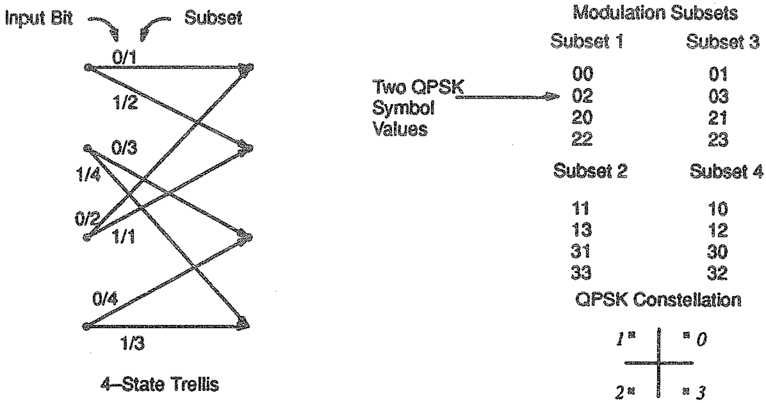

Figure 1 – Rate 3/4, 4–State, Trellis Coded Modulation

TRANSITION MATRICES

The LSPs have distribution and ordering characteristics that can be exploited when decoding. This can be done by looking at the LSP transitions, as each LSP value is dependent on each other LSP. The greatest dependency is between two adjacent LSPs. By measuring the joint probability distribution of two LSPs, 2–Dimensional transition matrices can be developed. An example of one is given in figure 2.

When constructing the trellis, the a–posteriori probabilities require the probability of a sequence of LSPs occurring. The length of the set of the LSPs increases as the trellis is constructed. To store all possible transition probabilities, of lengths up to 10 is unrealistic. Therefore to save on memory, the level of dependency is limited to two.

$$P(\omega_0..\omega_L) \approx \prod_{i=1}^{L} P(\omega_i|\omega_{i-1}) \, . \, P(\omega_0) \tag{1}$$

where ω_i = LSP i
and L is the length of the LSP sequence

Of course a better approximation would be to use transitions involving 3 LSPs but would require much more storage memory and involve a greater number of transition checks, increasing the computation. Knowledge of the LSP transition probabilities requires extensive training speech data. This speech data must try to be truly representative of all likely users of the system.

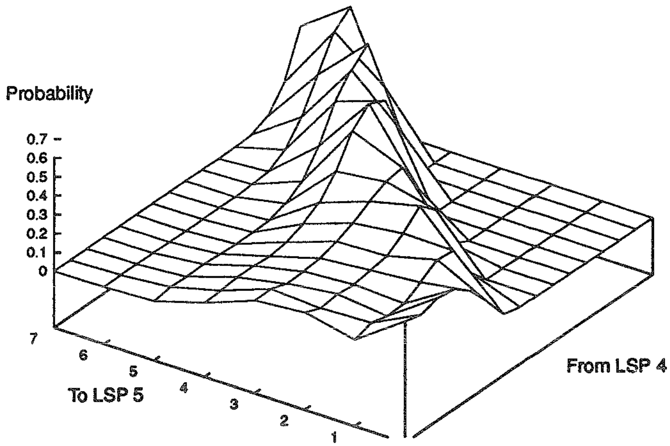

Figure 2 – LSP Transition Matrix from LSP 4 (16 levels) to LSP 5 (8 levels)

1st SCHEME – REMOVING ILLEGAL TRANSITIONS

This scheme is the simpler of the two presented. As the Viterbi decoder constructs the trellis, it uses the transition matrices to decide whether a branch in the trellis can occur. If the transition probability is zero then the corresponding branch in the trellis is disallowed. This provides greater accuracy when decoding the most probable path, by removing the possibility of illegal paths. To implement this system, the transition matrices do not need to store the transition probabilities but only whether a particular transition can or cannot occur. This reduces the amount of memory required. With this reduction, transitions greater than 2 could possibly be stored, however it is unsure how much gain is possible. The number of zero probability transitions in each 2–dimensional transition matrix varies, however the average is approximately 35%. This redundancy is enough to provide substantial BER improvement when using a Viterbi decoder. Table 2 shows the BER and spectral distortion (sd) results from using this scheme. Overall this scheme is very efficient (computation wise) as it reduces some of the computation in the Viterbi decoder and removes the need for a nonmonotonic correction algorithm (E.g. DoD LSP Replacement Algorithm [1]), as it guarantees that the LSPs are monotonic.

2nd SCHEME – USING MAP PROBABILITIES

A more accurate way of using the LSP probabilities in the Viterbi decoder is to combine them with the Euclidean distance modulation information to make up the branch metric. Required is the probability that a sequence of LSPs were sent, given a received sequence of symbols.
The a–posteriori probability is given as

$$P(\overline{\omega}|\overline{y}) = Pr\ (LSP\ sequence\ \overline{\omega}\ was\ sent\ given\ received\ symbol\ seq\ \overline{y}) \qquad (2)$$

Using Bayes Theorem

$$P(\overline{\omega}|\overline{y}) = \frac{P(\overline{y}|\overline{\omega})\ .\ P(\overline{\omega})}{P(\overline{y})} \qquad (3)$$

Considering the first numerator term, $P(\overline{y}|\overline{\omega})$ represents the effect the channel has on the received symbol sequence $\overline{y}$. Since the channel is independent for each symbol, equation (4) can be derived.

$$P(\bar{y}|\bar{\omega}) = \prod_{i=0}^{L} P(y_i|\omega_i) \tag{4}$$

For a Gaussian noise channel, $P(y_i|\omega_i)$ is the Gaussian pdf. Let σ_n^2 be the noise variance of the channel.

$$P(y_i|\omega_i) = \frac{1}{\sqrt{2\pi\sigma_n^2}} \; e^{-((y_{iI}-\omega_{iI})^2 + (y_{iQ}-\omega_{iQ})^2)/2\sigma_n^2} \tag{5}$$

where ω_{iI} and ω_{iQ} are the inphase and quadrature components of the modulated LSP ω_i.
The second numerator term of equation (3), $P(\bar{\omega})$, can be approximated by using the transition matrices presented in equation (1). To build up the trellis, we require a branch metric $B_{i,j}$ for each branch in the trellis. Each trellis branch corresponds to one LSP value.

$$B_{i,j} = P(\omega_{i,j}|y_i) \tag{6}$$

where $\omega_{i,j}$ = LSP number i, value index j.
Using (1), (3), (5) and (6), and removing all constants, the branch metric for the i[th] LSP, value index j can be reduced to

$$B_{i,j} = exp\left(\frac{-1}{2\sigma_n^2} \cdot \left[SED(y_i, \omega_{i,j})\right]\right) \cdot P(\omega_{i,j}|\omega_{i-1}) \tag{7}$$

where $SED(y_i, \omega_{i,j}) = (y_{iI} - \omega_{i,jI})^2 + (y_{iQ} - \omega_{i,jQ})^2$.

The state metric required in the Viterbi decoder is the a–posteriori probability of an LSP sequence given the received symbol sequence, i.e. $P(\bar{\omega}|\bar{y})$. This can easily be put in terms of branch metrics.

$$S_{k,l} = \prod_{i=0}^{k} B_{i,j} \tag{8}$$

where S_k = State metrics for LSP k.
$B_{i,j}$ = Branch metric for LSP i, value index j.

The state metric becomes the product of each previous chosen MAP branch metric, instead of the sum of SED's. Note however that the log of the branch metric can be taken for a reduction in computation. When building the trellis the number of dimensions of the transition matrices effects how much comparing is done. When two consecutive LSPs are tested, all possible values for both LSPs are compared. Therefore this involves varying and comparing all values of two consecutive branches in the trellis to find the best path. As a result, more computations are required per state compared with the standard Viterbi decoder. E.g. if 3–dimensional transition matrices are used then the decoder would have to look back 3 branches to find the most probable path to that state.

RESULTS

The measure used to compare the different schemes is the standard LSP measure – spectral distortion (sd) [2]. The transition matrices were trained using 2 minutes of the speech taken from the Timit speech database, using 32 different speakers. Results are presented in tables 1 – 3 and figure 3. Notice that for both schemes, tables 2 and 3, the BER performance for the low BER channels is worse than the standard system. This is due to a slight mismatch in the transition matrices to the test speech data. However this has only a minor effect on the output speech quality and goes largely unnoticed. At high BER channels, such as 10% BER, both schemes clearly outperform the standard system, with the MAP scheme handling the errors the best. Figure 3 also shows the sd plot of standard QPSK modulation using the Dod replacement algorithm for reference. Of the objective measures below, often an important measure is the percentage of sd outliers > 4 dB. When the sd is greater than 4 dB it usually results in noticable distortion. Therefore it is an important measure for the high BER channels. From this measure, the performance gain from the two schemes can really be noticed, with the MAP scheme performing the best.

A bursty error channel, such as Mobilesat™ [8], if in heavy shadowing can produce BER's of up to 20%. This is where the use of MAP probabilities in the Viterbi decoder can remove much of the distortion due to an error burst. Through listening tests, the MAP LSP decoder, outperforms both schemes. It manages to remove the large bursts, reducing them to only small distortions. The 1st scheme will decrease the amount of distortion during bursts, but it doesn't perform as well as the MAP LSP decoder. Since the MAP LSP decoder only has real performance gains for very high BER, it is therefore highly suited to bursty error channels. For a channel which rarely goes higher than 5 % BER the 1st scheme would probably be the better choice due to the lower memory and computation requirements. Therefore a choice of which scheme to use would be dependent highly on the channel characteristics.

Table 1 – Rate 3/4 TCM with DoD LSP Replacement Algorithm

BER for QPSK	SD (dB)	% Outliers 2–4dB	% Outliers >4dB	BER
0.0	1.477190	9.45	0.0	0.0
0.001	1.478254	9.51	0.0	6.356×10^{-5}
0.01	1.513404	10.97	0.27	0.001811
0.02	1.574006	13.19	0.81	0.004322
0.05	2.088064	23.12	7.83	0.0234
0.1	4.083836	29.12	40.19	0.1023
0.2	9.257094	6.75	91.36	0.2663

Table 2 – Rate 3/4 TCM with LSP Restrictions

BER for QPSK	SD (dB)	% Outliers 2–4dB	% Outliers >4dB	BER
0.0	1.487758	9.94	0.05	3.496×10^{-4}
0.001	1.488345	10.16	0.0	3.972×10^{-4}
0.01	1.513565	11.72	0.0	0.001764
0.02	1.558552	13.88	0.32	0.003893
0.05	1.865503	23.77	4.05	0.01529
0.1	2.99744	36.09	21.02	0.05290
0.2	6.321909	24.10	70.34	0.1821

Table 3 – Rate 3/4 TCM with LSP MAP Probabilities (σ_n^2 set to 0.05 BER)

BER for QPSK	SD (dB)	% Outliers 2–4dB	% Outliers >4dB	BER
0.0	1.489582	10.05	0.11	3.496×10^{-4}
0.001	1.509009	11.08	0.22	0.00135–62
0.01	1.543701	12.8	0.38	0.003209
0.02	1.570444	14.26	0.38	0.004592
0.05	1.775222	20.85	2.86	0.01374
0.1	2.477194	32.9	13.29	0.04146
0.2	5.475156	32.04	58.29	0.1590

For tables 1 to 3, the BER is quite high due to the QPSK errors of the 4 bits and also due to the tail of the trellis causing large number of bit errors. The tail of the trellis could be extended by adding lower sensitive bits, which would improve both BER and sd results, however the last couple of LSPs are less sensitive than the others. Varying the expected noise variance σ_n^2 gives different results, as it is a weighting factor between the two probabilities. For table 3, the variance was set to QPSK BER of 0.05, providing optimum performance for that particular channel.

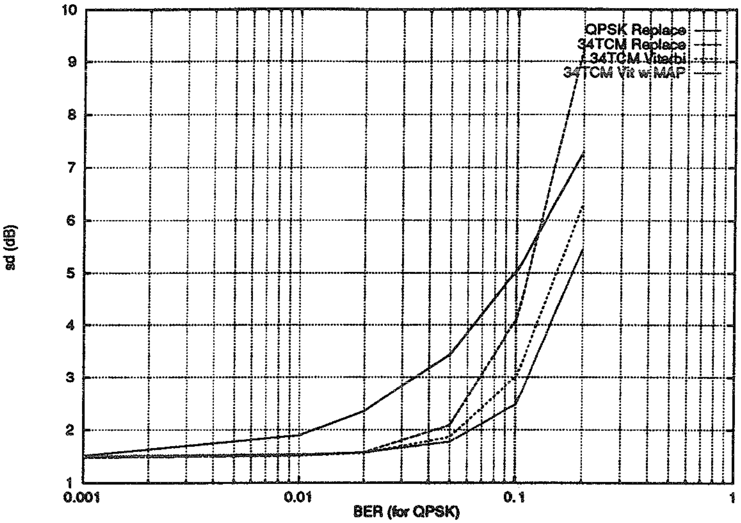

Figure 3 – Comparison of the three different decoders, including QPSK with replacement

CONCLUSION

From the results it is noted that the MAP LSP Decoder outperforms the other two systems. This must be weighted against the extra computation when decoding. For many of the static BER channels, the 1st scheme can perform just as well as the MAP scheme, and has much less computation. Note that both new schemes ensure filter stability by disallowing zero probability transitions. This is compared to the standard system which must use some post algorithm (e.g. LSP Replacement Algorithm [1]) to maintain monotonic LSP values. The two schemes presented are designed to work for very poor channels, and thus are well suited to bursty error channels. Also the two schemes developed have the ability to be switched on and off if the channel information is available. This would be particularly useful for a bursty error channel. This method of using the residual redundancy of the scalar quantisation of the LSPs can be used with and without a channel coder. Alternative channel coding could be employed depending on the channel characteristics. However there is a distinct advantage in using a convolutional coder as it requires a Viterbi decoder.

REFERENCES

[1] National Communications System, "Details to Assist in Implementation of Federal Standard 1016 CELP", Jan. 1992.

[2] P. Secker & A. Perkis, "Joint source and channel trellis coding of line spectrum Pair Parameters", Speech Communication II, pp 149–158, June 1992.

[3] K.K. Paliwal & B.S. Atal, "Efficient Vector Quantization of LPC Parameters at 24 Bits/Frame", IEEE Transactions on Speech and Audio Processing, vol. 1, no. 1, Jan 1993.

[4] N. Phamdo & N. Farvardin, "Optimal Detection of Discrete Markov Sources Over Discrete Memoryless Channels – Applications to Combined Source-Channel Coding", IEEE Transations on Information Theory, vol. 40, no. 1, Jan 94

[5] D. Divsalar & M.K. Simon, "Multiple Trellis Coded Modulation (MTCM)", IEEE Trans. on Communications, vol. 36, pp. 410–418, Apr. 1988.

[6] S.S. Pietrobon, et al. "Trellis-Coded Multidimensional Phase Modulation", IEEE Transactions on Information Theory, Vol. 36, No.1, Jan 1990.

[7] W.N. Farrell & W.G. Cowley, "A Comparative Study of Modulation Schemes for the Transmission of CELP Parameters", Proc. DSP for Comm Systems, April 1994, Adelaide

[8] K. Dinh, N. Hart, I.S. Jorgensen, "mobilsat^{TM} – Specifications and Performance", Proceedings IREECON'91, pp 166–169.