Results: First, computer simulations with a 7-chip *m*-sequence were performed to prove the applicability of the analysis. In the simulations 40000 code periods were used in equal power two-path channels and the complex gains of both paths were randomly changed every 50 code periods. When comparing the theoretical and simulated P_ds and P_{fa}s, as well as T_{MA}s as functions of signal-to-noise ratio E_b/N_0, no significant difference in the results was seen thus proving the applicability. Fig. 2 presents the theoretical T_{MA}s when the P_{fa} per cell is 10^{-2} when the code length is a parameter (N = 3, 7 and 15). The false alarm penalty time is $100T_b$. The increase in T_{MA} due to non-ideal ACF can be calculated as $\Delta_{MA} = 100(T_{MA1} - T_{MA2})/T_{MA2}$, where T_{MA1} and T_{MA2} are the mean acquisition times obtained by using real spreading codes (m-sequence) with non-ideal ACF, and with ideal ACF, respectively. The increase in the worst case is now ~100% for N = 3, 20% for N = 7 and < 10% for N = 15. Fig. 3 presents the increase in the case when the P_{fa} per cell is 10^{-4}. The increase in T_{MA} can be significant due to IPI, especially when low P_{fa} is required.

Conclusions: Multipath propagation causes intersymbol interference in conventional narrowband communication systems and interpath interference in spread spectrum communication systems. In this Letter we have considered the influence of interpath interference (IPI) on the code acquisition performance. The acquisition was performed by a matched filter. The results indicate that with short spreading codes the effect of interpath interference can be significant. It should be taken into account that *m*-sequences have 'ideal-like' autocorrelation functions. Their out-of-phase correlation is –1, and, for example, using Gold sequences, the effect of interpath interference can be assumed to be more significant. To obtain exact results, numerous simulations should be performed, due to the non-smooth ACF of Gold codes.

Acknowledgments: This work was supported by the Technology Development Centre of Finland, Nokia, and the Finnish Air Force.

6 April 1999
Electronics Letters Online No: 19990664
DOI: 10.1049/el:19990664

J. Iinatti (*University of Oulu, Telecommunication Laboratory Centre for Wireless Communication, PO Box 4500, FIN-90401 Oulu, Finland*)

E-mail: jari.iinatti@ee.oulu.fi

S. Kerhuel (*University of Rennes, 20 Avenue des buttes de coësmes, Residence A, 35043 Rennes Cedex, France*)

References

1 POLYDOROS, A.: 'On the synchronization aspects in direct-sequence spread spectrum systems'. PhD. Dissertation, University of Southern California, Los Angeles, CA, USA, 1982
2 IINATTI, J.: 'Matched filter code acquisition employing a median filter in direct sequence spread-spectrum systems with jamming'. Acta Universitatis Ouluensis C 102, Doctoral Thesis, University of Oulu, Oulu, Finland, 1997
3 IINATTI, J.: 'Mean acquisition time of DS code synchronization in fixed multipath channel'. Proc. ISSSTA'98, Sun City, South Africa, 2–3 September 1998, Vol. 1, pp. 116–120

Efficient quantisation method for LSF parameters based on restricted temporal decomposition

Sung-Joo Kim and Yung-Hwan Oh

A restricted temporal decomposition method is presented for line spectral frequency (LSF) parameters. The proposed method interpolates an LSF vector trajectory efficiently while conserving the LSF ordering property. Experimental results show that interpolated LSF parameters can be transparently quantised at a rate of 753bit/s.

Introduction: Temporal decomposition is a speech coding method which decomposes a given vector trajectory into a set of temporally overlapped event functions and corresponding event vectors. The original temporal decomposition estimates events with some restrictions on event functions such as time localisation, but does not consider any restriction on event vectors. This method has been successfully applied to quantising log-area, log-area-ratio, or cepstrum parameters [1 – 3]. However, when we apply it to decomposing LSF parameters holding an ordering property [4], a serious problem occurs. Some event vectors are outside of the valid range of LSF parameters or disordered so that they cannot be interpreted as LSF parameters. Consequently, these event vectors do not have corresponding stable spectra and cannot be quantised effectively as LSF parameters. To solve this problem, we propose a restricted temporal decomposition (RTD) method so that every event vector for an LSF vector trajectory preserves the property of LSF parameters and that the vector trajectory can be interpreted as the interpolation of the estimated events.

Restricted temporal decomposition: Let a given vector trajectory and its corresponding sets of event vectors and functions be $Y = [\vec{y}(1), \vec{y}(2), ..., \vec{y}(N)]$, $A = [\vec{a}_1, \vec{a}_2, ..., \vec{a}_J]$, and $\Phi = [\vec{\phi}_1, \vec{\phi}_2, ..., \vec{\phi}_J]^T$, respectively, where $\vec{y}(n) = [y_1(n), y_2(n), ..., y_p(n)]^T$ is the LSF vector sampled at time n, $\vec{a}_j\ [a_{j,1}, a_{j,2}, ..., a_{j,p}]^T$ the jth event vector, and $\vec{\phi}_j = [\phi_j(1), \phi_j(2), ..., \phi_j(N)]^T$ the jth event function. The temporal decomposition estimates the event vectors and functions which minimise

$$E = \|Y - A\Phi\|^2 = \sum_{n=1}^{N} \|\vec{y}(n) - \vec{y}'(n)\|^2 \qquad (1)$$

where

$$\vec{y}'(n) = \sum_{j=1}^{J} \vec{a}_j \phi_j(n) \qquad (2)$$

$\phi_j(n)$ is restricted to the range [0, 1] and to having maximum value at its corresponding centre $C(j)$. Events are ordered with respect to their central positions. For RTD, we propose the use of an interpolation, eqn. 3 instead of eqn. 2, assuming that the jth event function $\phi_j(n)$ has a nonzero value only for $C(j-1) < n < C(j+1)$ and that the sum of all event functions is one at any time n.

$$\vec{y}'(n) = \vec{a}_j\phi_j(n) + \vec{a}_{j+1}\phi_{j+1}(n) = \vec{a}_j\phi_j(n) + \vec{a}_{j+1}(1-\phi_j(n)) \quad \text{for } C(j) \le n < C(j+1) \qquad (3)$$

Substituting eqn. 3 into eqn. 1, we obtain

$$E = \sum_{j=1}^{J-1} \sum_{n=C(j)}^{C(j+1)-1} \left\| (\vec{y}(n) - \vec{a}_{j+1}) - (\vec{a}_j - \vec{a}_{j+1})\phi_j(n) \right\|^2 \qquad (4)$$

where $C(1) = 1$, $C(J) = N + 1$.

To minimise the error E, $\phi_j(n)$ should be estimated using eqn. 5, which is obtained from setting the partial derivatives of eqn. 4 with respect to $\phi_j(n)$ equal to zero:

$$\hat{\phi}_j(n) = \frac{\langle (\vec{y}(n) - \vec{a}_{j+1}), (\vec{a}_j - \vec{a}_{j+1}) \rangle}{\left\| \vec{a}_j - \vec{a}_{j+1} \right\|^2} \qquad (5)$$

Finally, considering the above restrictions on event functions, we determine $\phi_j(n)$ as

$$\phi_j(n) = \begin{cases} 1 - \phi_{j-1}(n) & \text{if } C(j-1) < n < C(j) \\ 1 & \text{if } n = C(j) \\ \min(1, \max(0, \hat{\phi}_j(n))) & \text{if } C(j) < n < C(j+1) \\ 0 & \text{otherwise} \end{cases} \qquad (6)$$

Turning to the event vectors, we estimate them to correspond to the determined event functions in the minimum squared error sense by using the following formula [1 – 3]. Note that $(\Phi\Phi^T)$ here is tridiagonal and easily inverted:

$$A = Y\Phi^T(\Phi\Phi^T)^{-1} \qquad (7)$$

However, the estimated event vectors may violate the ordering property since the error criterion does not consider this property, so if $\vec{a}_j$ with $a_{j,k-1} \ge a_{j,k}$ minimises error E, then $a_{j,k-1}$ and $a_{j,k}$ should be changed to $a'_{j,k-1}$ and $a'_{j,k} = a'_{j,k-1} + \varepsilon$. Considering that the increment of error E caused by this change is

$$\Delta = \sum_n (a_{j,k-1} - a'_{j,k-1})^2 \phi_j(n)^2 + \sum_n (a_{j,k} - a'_{j,k})^2 \phi_j(n)^2 \qquad (8)$$

$a'_{j,k-1}$ should be set as follows to minimise Δ. Here, we use ε = 0.01 empirically:

$$a'_{j,k-1} = (a_{j,k-1} + a_{j,k} - \varepsilon)/2 \qquad (9)$$

In the result, when the central positions of the events $C(j)$, $j = 1, ..., J$, are known, and the corresponding event vectors are initialised with the samples of the vector trajectory $\vec{y}(C(j))$, we can calculate proper event functions and vectors iteratively by using eqns. 6 and 7. We suggest using the local minimal points of the following spectral transition measure based on LSF parameters as the central positions of corresponding events:

$$STM_{LSF}(n) = \left\| \sum_{t=-M}^{M} t \cdot \vec{y}(n+t) \right\|^2 \qquad (10)$$

where $M = 2$. However, we have experimentally determined that the number of events found with $STM_{LSF}(n)$ is ~ 10 events per second and not enough for interpolating an LSF vector trajectory. Therefore, we insert a new event where the initial interpolation error $e(n) = \|\vec{y}(n) - \vec{y}'(n)\|^2$ has a local maximum and is larger than the certain threshold θ. For online analysis, we segment the input vector trajectory and perform the RTD of each segment sequentially. A segment can be bounded by two event centres found with $STM_{LSF}(n)$. Since an event function is overlapped by its adjacents, the last event of each segment should be re-estimated with the following segment. The whole RTD algorithm is summarised as follows.

Step 1: initialise $C(1) \leftarrow 1$, $\vec{a}_1 \leftarrow \vec{y}(1)$, and $J2$

Step 2: find $C(J)$, the next local minimum point of $STM_{LSF}(n)$, which will be the end of the current segment; set $\vec{a}_J \leftarrow \vec{y}(C(J))$

Step 3: estimate the initial event functions $\vec{\phi}_j$ corresponding to the current set of event vectors using eqn. 6; note that no iteration is needed for this step

Step 4: if $\max_n (e(n)) > \theta$, insert an event as $J \leftarrow J + 1$, $C(J) \leftarrow \arg\max_n (e(n))$, $\vec{a}_J \leftarrow \vec{y}(C(J))$, and reorder the events by their central positions, then go back to step 3

Step 5: re-estimate the event vectors $\vec{a}_j$ using eqn. 7, but do not update $\vec{a}_1$ if it is from the previous segment

Step 6: re-estimate the event functions using eqn. 6; if the results have converged or have been re-estimated a certain number of times, go to step 7; if not, go back to step 5

Step 7: store the events for the current segment; to analyse the next segment, set $C(1) \leftarrow C(J-1)$, $C(2) \leftarrow C(J)$, $\vec{a}_1 \leftarrow \vec{a}_{J-1}$, $\vec{a}_2 \leftarrow \vec{a}_J$, and $J \leftarrow 3$, then go back to step 2

Experiments: We have designed two separate experiments to measure the performance of the interpolation and quantisation based on RTD. We used prediction gain to measure the interpolation performance and spectral distortion (SD) to measure the quantisation performance [5]. During these experiments, we used the weighted squared Euclidean distortion of the LSF parameters as described in [4] and we derived eqns. 6, 7 and 9 based on it. As the speech corpus for these experiments, we chose 1890 phonetically-diverse sentences (SI set) from the TIMIT database and used 504 sentences for testing and 1386 for training. A 10th order LPC analysis was performed using the autocorrelation method with a 30ms Hamming window which was shifted every 20ms. Finally, the LPC parameters were converted into the LSF parameters.

For the 221586 frame long training set, the average prediction gain of the original LSF parameters was 9.09dB and the gain reduction caused by interpolation was only 0.15dB when we used θ = 0.6. In this case, the average event occurrence frequency was 18.16Hz. An event function $\phi_j(n)$ can be represented by its length $p(j) = C(j+1) - C(j)$ and shape in $[C(j), C(j+1)]$. The maximum event length was 11 frames, so we coded $p(j)$ with four bits except for $p(j) = 1$ with three bits, '000'. For quantising the shape, we first took 10 equidistant samples of the event function by interpolation and then vector-quantised them. A vector quantiser for the shape of an event function and 10 scalar quantisers for the event vector dLSFs were trained using the interpolation results of the training set.

Finally, we quantised the events obtained from the RTD of the 83910 frame long test set, varying the bit allocations. We measured the average SD between the interpolated LSF parameters and the quantised parameters. Table 1 shows the interpolated LSF parameters are quantised transparently [5] when 33 bits are used for an event vector and 6 bits for the shape of an event function. Although the overall SD caused by both interpolation and quantisation was 1.74dB and could not satisfy the conditions for transparent coding, most reconstructed speech sentences could not be easily distinguished from the original sentences during the informal listening test. Table 2 shows average bit rates for this configuration.

Table 1: Performance of quantiser for interpolated LSF vector trajectory

Bit allocation		SD	Outliers	
dLSF SQ	φ shape VQ		2–4dB	> 4dB
		dB	%	%
31 bit	4 bit	1.126	4.89	0.16
(3,3,3,3,4,	5 bit	1.070	3.97	0.16
3,3,3,3,3)	6 bit	1.011	3.25	0.15
32 bit	4 bit	1.082	3.60	0.08
(3,3,3,3,4,	5 bit	1.023	2.86	0.07
3,4,3,3,3)	6 bit	0.963	2.17	0.07
33 bit	4 bit	1.054	2.92	0.04
(3,3,3,3,4,	5 bit	0.994	2.20	0.04
3,4,3,4,3)	6 bit	0.933	1.57	0.03

Table 2: Bit rate of proposed quantiser

	dLSF SQ	Event function		Frequency	Total
		Position	Shape		
				Hz	bit/s
$p(j) > 1$	33	4	6	14.04	604
$p(j) = 1$	33	3	0	4.12	149
Total	753 bit/s				

Conclusion and discussion: We have performed a restricted temporal decomposition of the LSF parameters and presented a quantisation method for the resulting events. As a result, we were able to interpolate an LSF vector trajectory with ~ 18 events/s while the loss of prediction gain was only 0.15dB. We were able to transparently quantise the events at 753bit/s. The bit rate may be further reduced if we use a vector quantiser for event vectors. While the overall distortion could not satisfy the condition for transparent coding, the quality of reconstructed speech is sufficient for a low-bit-rate vocoder. The coding delay of this method was nonuniform because the event length varied. The average event length was 55ms and its standard deviation was 29.45ms while interpolating the training set. The proposed method reduces the computational complexity of the temporal decomposition methods described in [1 – 3] because it does not make use of singular value decomposition nor adaptive Gauss-Seidel iterations in estimating event functions. Moreover, the error decreases monotonically and converges rapidly to a local minimum since eqns. 6, 7 and 9 lead to stepwise optimal solutions, so re-estimating five times is sufficient for for the interpolation error to converge.

The proposed method differs from straightforward interpolation methods in several ways: in the re-estimation of event vectors, in the iterative optimisation, and in the adaptive selection of update points. In fact, these differences are taken from the temporal decomposition method. Furthermore, the estimated events also have desirable properties such as time localisation and correspondence to speech events, which the temporal decomposition aims for. Hence, the proposed method can be considered as a reasonable adaptation of temporal decomposition for LSF parameters.

22 March 1999
Electronics Letters Online No: 19990670
DOI: 10.1049/el:19990670

Sung-Joo Kim and Yung-Hwan Oh (*Department of Computer Science, Korea Advanced Institute of Science and Technology, 373-1 Kusong-dong, Yusong-gu, Taejon 305-701, Korea*)

References

1 ATAL, B.S.: 'Efficient coding of LPC parameters by temporal decomposition'. Int. Conf. Acoustics, Speech and Signal Processing, 1983, pp. 81–84

2 VAN DIJK-KAPPERS, A.M.L., and MARCUS, S.M.: 'Temporal decomposition of speech', *Speech Commun.*, 1989, **8**, (2), pp. 125–135

3 CHENG, Y.-M., and O'SHAUGHNESSY, D.: 'On 450-600b/s natural sounding speech coding', *IEEE Trans.*, 1993, **ASSP-1**, (2), pp. 207–220

4 RAMACHANDRAN, R.P., *et al.*: 'A two codebook format for robust quantization of line spectral frequencies', *IEEE Trans. Speech Audio Process.*, 1995, **3**, (3), pp. 157–168

5 LEMMA, A.N., *et al.*: 'LPC quantization using wavelet based temporal decomposition of the LSF'. EUROSPEECH'97, 1997, pp. 1259–1262

Evaluation of extinction ratio induced performance penalty due to interferometric noise

J.E. Mitchell, P.M. Lane and J.J. O'Reilly

A rigorous assessment of the impact of finite extinction ratio in systems corrupted with interferometric noise is presented. In contrast to previous work it is shown that the extinction ratio has minimal effect on the overall crosstalk tolerance, with signal-crosstalk beating between data '1' terms being dominant.

Introduction: Interferometric crosstalk has long been considered a major factor in the overall performance limit of multiwavelength optical networks. Interferometric noise (IN) is formed in such systems owing to imperfect isolation in routing and demultiplexing components, allowing other signals of nominally the same wavelength to fall on the receiver photodiode with the desired signal. The interaction of these signals causes the generation of beat components that cannot be characterised using the same methods applied to traditional crosstalk power. In recent years a number of advances have been made in the accuracy of the statistical models used to describe this phenomenon; however, to date, most assessments have used variations on Gaussian approximation (GA) to evaluate this effect. It has been reported on numerous occasions that the use of Gaussian statistics is not well suited to the study of IN unless the system is such that the central limit theorem can be relied on, i.e. that there is a large number of interfering terms [1, 2]. The effect of finite extinction ratio is present in virtually every real system; however, assessments of its contribution to any system penalty have only made use of Gaussian statistics without fully considering the obvious impact of symbol conditioning [3].

A simple form of GA evaluates the power penalty by following a procedure similar to that used in [4], although it may be modified for finite extinction ratio, as demonstrated in [5]. In this Letter we assess IN by considering the actual composite statistics of all noise contributions, including individually the effect of each beat term due to the different symbol combinations. To effect this, we use the modified Chernoff bound (MCB) method [6] in view of its ease of formulation and computational efficiency.

Interferometric beat noise: The signal-crosstalk beat noise component at the receiver can be described as

$$i = 2P_s \sum_{i=1}^{N} m_s(t) m_i(t) \sqrt{\epsilon_i} \cos[(\omega_s - \omega_i)t + \phi_s(t) - \phi_i(t)] \qquad (1)$$

where P_s is the optical signal power, ϵ_i the relative crosstalk power ($= P_i/P_s$), s and i as subscripts denote expected and interfering signal, respectively, ω is the optical frequency, $\phi(t)$ the optical phase, N the number of interferers and $m(t)$ represents the binary symbols forming the message: $m(t) \in \{r, 1\}(0 < r < 1)$ with r accounting for the extinction ratio factor. If we consider the wavelengths to be nominally identical, the most likely scenario in a dense WDM system, we see that the term inside the summation in eqn. 1 is of the form $\alpha\cos(\phi)$, the probability density function (PDF) of which is the well-known arc-sinusoidal function. From this the moment generating function (MGF) is readily obtained since the addition of independent random variables corresponds to the multiplication of MGFs [7]:

$$M_{IN}(s) = \prod_{i=1}^{N} I_0(\alpha s) \qquad (2)$$

where $I_0(\bullet)$ is the modified Bessel function, first kind, zero-order. Laser biasing or modulator extinction constraints will inevitably lead to a finite extinction ratio in practical optical systems. 'Traditional' evaluation methods consider this a simple degradation of the eye opening by including a factor of $(1 - r)/(1 + r)$ in the Q value, while systems corrupted by IN will, in addition, have terms owing to the signal-crosstalk beating. In a model with perfect extinction in all components it is assumed that a datum '0' is symbolised by the complete absence of light, which means that no beating will occur if either the signal or crosstalk has a datum of '0'. Now we must consider that a datum '0' is represented by a signal power of rP_s, allowing the beating between the signal and each interferer to be governed by one of four possible r.v.s dependent on the symbol conditions.

Impact of finite extinction ratio: Here we compare two evaluation methods: the Gaussian approximation (GA) method (decision threshold non-optimised) and the modified Chernoff bound (MCB) method, defined as eqns 3 and 4, respectively:

$$BER \simeq \frac{1}{2} Q\left[\frac{i_p - D}{\sqrt{\sigma_n^2 + \sigma_{IN1}^2}}\right] + \frac{1}{2} Q\left[\frac{D - ri_p}{\sqrt{\sigma_n^2 + \sigma_{IN0}^2}}\right] \qquad (3)$$

$$P_e \leq \frac{M_G(s)}{2s\sigma_n\sqrt{2\pi}}\left[M_{I(0)}(s)e^{(-sD)} + M_{I(1)}(s)e^{-s(i_p - D)}\right] \quad s > 0 \qquad (4)$$

where D represents the decision threshold, σ_n^2 is the thermal noise variance, with σ_{IN1}^2 and σ_{IN0}^2 the beat noise variance for signals with values of '1' and '0', respectively. The performance degradation is evaluated at various extinction ratios including –8dB (ITU-T Recommendation G.957(7/95)). Previous work using GA with an optimised decision threshold has suggested that the inclusion of a finite extinction ratio affects the maximum tolerable crosstalk level [3].

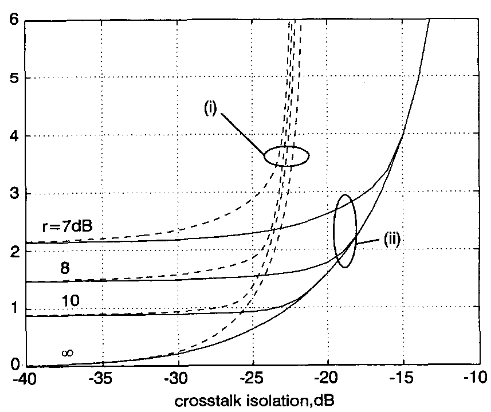

Fig. 1 *Power penalty against crosstalk isolation calculated using Gaussian approximation and modified Chernoff bound methods*

(i) Gaussian approximation
(ii) rigorous evaluation

The calculation using the MCD method, using an average power decision threshold, as shown in Fig. 1, demonstrates that finite extinction ratio inherently causes eye closure, as can be seen by the marked power penalty floor in keeping with establish theory [8]. However, we find now that there is a single crosstalk limit regardless of the extinction ratio. From this we deduce that even though additional beating terms are introduced by the finite extinction ratio, the beating of the two data values of '1' still dom-