

Saikat Chatterjee and T.V.Sreenivas

Department of Electrical Communication Engineering Indian Institute of Science, Bangalore, India

saikat@ece.iisc.ernet.in, tvsree@ece.iisc.ernet.in

Abstract

We investigate the use of a two stage transform vector quantizer (TSTVQ) for coding of line spectral frequency (LSF) parameters in wideband speech coding. The first stage quantizer of TSTVQ, provides better matching of source distribution and the second stage quantizer provides additional coding gain through using an individual cluster specific decorrelating transform and variance normalization. Further coding gain is shown to be achieved by exploiting the slow time-varying nature of speech spectra and thus using inter-frame cluster continuity (ICC) property in the first stage of TSTVQ method. The proposed method saves 3-4 bits and reduces the computational complexity by 58-66%, compared to the traditional split vector quantizer (SVQ), but at the expense of 1.5-2.5 times of memory.

Index Terms : Vector Quantization, LSF Coding.

1. Introduction

Most of the wide-band speech coders use linear prediction (LP) analysis of 16-th order and thus effective vector quantizer (VQ) of the LP coefficients (LPCs), in terms of LSFs, is in great demand.

Effective VQ of LSFs involves several performance measures which are inter-related and worth investigating; the different measures are the coding delay, bit-rate, objective/perceptual distortion, computational and memory complexity. Due to the enormous complexity associated with an unconstrained full search VQ, various forms of sub-optimal VQs, such as tree-search [1], multistage VQ (MSVQ) [4], are proposed for telephone-band speech and further extended to wideband speech coding ([5], [9]). The most cited, successful and practically used technique is split VQ (SVQ) [3]; SVQ reduces the computational and memory complexity by designing and operating VQs of smaller dimension. However, the performance would be poorer than a full search VQ due to independent treatment of the sub-vectors. Lefebvre et al. [6] and Chen et al. [7] used a seven part SVQ operating at 49 bits/frame to quantize LSF of wideband speech.

Among the recent techniques, a split-multistage VQ (S-MSVQ) with MA predictor, is used to quantize the LPC parameters in AMR-WB speech coder [12]. A multi-frame GMM based block quantizer is proposed in [13] which provides good performance at the expense of higher coding delay. A scheme of switched SVQ (SSVQ) is proposed in [14] which saves upto 4 bits and reduces the computational complexity by 24-41%, compared to the five split SVQ method, but requires 7-11 times of memory.

In this paper, we investigate a two stage VQ along with SVQ for LSF quantization in wideband speech. Inclusion of a decorrelating transform, at the second stage, improves the quantizer performance. We refer to this new method as two stage transform VQ (TSTVQ). Further, we show that the slow time-varying nature of speech spectra can be exploited at the first stage and higher coding gain can be achieved using ICC property. The proposed TSTVQ provides better rate-distortion performance than the traditional SVQ even at a much lower computational complexity.

2. Two stage transform VQ

The goal of most of the structured VQs is to save on complexity such that an acceptable performance is realized at a bit-rate somewhat higher than a full search VQ. In many mobile speech coding applications, both for wideband and low bit rate telephony speech, it is important to reduce computation as well as memory. From the VQ literature [2], we know that MSVQ provides both the advantages of reduced memory and reduced complexity, unlike treestructured VQ. Therefore, we explore the use of a two stage VQ method. Fig. 1 (a) shows the the block diagram of a two stage VQ, in which the input vector, \mathbf{X} , is first quantized to \mathbf{X}_1 using the first stage quantizer Q_1 and the residual vector, \mathbf{E}_2 , is quantized to $\hat{\mathbf{E}}_2$ using the second stage quantizer Q_2 ; at the receiver, the reproduction vector is realized as: $\hat{\mathbf{X}} = \hat{\mathbf{X}}_1 + \hat{\mathbf{E}}_2$. The first stage quantizer, Q_1 , has b_0 bits allocated, by which M number of Voronoi regions $(M = 2^{b_0})$ are formed in the original vector space, whose centroids are the reconstruction vectors, $\{\hat{\mathbf{X}}_1\}$. Unlike a basic MSVQ scheme in which the residue vector, $\mathbf{E}_2 = \mathbf{X} - \hat{\mathbf{X}}_1$, is coded directly using Q_2 , we perform a decorrelation and variance normalization of the residue vector before applying to Q_2 . Therefore, residue vectors of all the Voronoi regions get decorrelated and have unit variance along all dimensions. This will permit Q_2 to be implemented using a bunch of scalar quantizers or using an SVQ technique to achieve some higher dimensional coding advantages. We use SVQ technique to realize Q_2 and thus it is important to use an appropriate distance measure to mitigate the effect of variance normalization.

2.1. TSTVQ coding

A block diagram of TSTVQ is shown in Fig. 1 (b), consisting of three parts: Q_1 , transform block and a common SVQ quantizer. The *M* Voronoi regions, at first stage quantizer, are referred to as clusters and thus the *M* code-vectors of Q_1 , are the cluster mean vectors, $\{\mu^k\}_{k=1}^M$, using simple Euclidean distance (ED) measure.

Let, \mathbf{x}^i be the *i*th frame 16-dimensional LSF vector, which is quantized to μ^k using Q_1 and thus belongs to the *k*th cluster. Then, the transformed and normalized residual vector is:

$$\mathbf{u}^{i} = \Lambda^{k} \operatorname{T}^{k} \left[\mathbf{x}^{i} - \boldsymbol{\mu}^{k} \right]$$
 (1)



Figure 1: (a) Two stage VQ. (b) Two stage transform VQ (TSTVQ).

where μ^k is the mean vector and \mathbf{T}^k is the KLT matrix for the kth cluster; $\Lambda^k = diag\left[\left\{1/\sqrt{\lambda_j^k}\right\}_{j=1}^{16}\right]$, where $\{\lambda_j^k\}_{j=1}^{16}$ are the eigen values of the covariance matrix of kth cluster. Thus, the indices of the cluster mean, μ^k , of first stage and the coded vector $\hat{\mathbf{u}}^i$ of second stage have to transmitted over the channel. The coding gain of the kth cluster specific KLT, \mathbf{T}^k , is higher as it is tuned to the source distribution of the kth cluster. Due to the different covariance structures of the clusters, the variance normalization of the transformed residual vector is carried out to achieve a common data spread (of unity variance) of all the clusters and thus paves the way for efficient use of a common SVQ codebook in Q_2 .

2.2. Cluster specific weighted Euclidean distance

It is common to use weighted Euclidean distance (WED) measure to search the VQ codebook in the context of LSF quantization ([3], [8]). For the *i*th frame, WED, between the input vector (\mathbf{x}^i) and the coded vector $(\hat{\mathbf{x}}^i)$, is defined as:

$$d(\mathbf{x}^{i}, \hat{\mathbf{x}}^{i}) = (\mathbf{x}^{i} - \hat{\mathbf{x}}^{i})^{T} W^{i}(\mathbf{x}^{i} - \hat{\mathbf{x}}^{i}) = \sum_{j=1}^{16} w_{j}^{i} (x_{j}^{i} - \hat{x}_{j}^{i})^{2} \quad (2)$$

where W^i is a diagonal weighting matrix with elements as $\{w_j^i\}_{j=1}^{16}$, which is dependent on the *i*th LSF vector. From Eqn. 1, we observe that the original vector, $\mathbf{x}^i = [\Lambda^k \mathbf{T}^k]^{-1} \mathbf{u}^i + \mu^k$; thus the decoded vector at the receiver is realized as: $\hat{\mathbf{x}}^i = [\Lambda^k \mathbf{T}^k]^{-1} \hat{\mathbf{u}}^i + \mu^k$. Therefore, simplifying the WED measure, given in Eqn. 2, in terms of original and coded transformed coefficients (\mathbf{u}^i and $\hat{\mathbf{u}}^i$) at second stage, we can write:

$$d(\mathbf{x}^{i}, \hat{\mathbf{x}}^{i}) = \begin{bmatrix} \mathbf{u}^{i} - \hat{\mathbf{u}}^{i} \end{bmatrix}^{T} \begin{bmatrix} [\Lambda^{k} \mathbf{T}^{k}]^{-1} \end{bmatrix}^{T} W^{i} \begin{bmatrix} \Lambda^{k} \mathbf{T}^{k} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{u}^{i} - \hat{\mathbf{u}}^{i} \end{bmatrix}$$
$$= \begin{bmatrix} \mathbf{u}^{i} - \hat{\mathbf{u}}^{i} \end{bmatrix}^{T} O^{k,i} \begin{bmatrix} \mathbf{u}^{i} - \hat{\mathbf{u}}^{i} \end{bmatrix}$$
(3)

In Eqn. 3, the new weighting matrix, denoted by $O^{k,i} = [[\Lambda^k T^k]^{-1}]^T W^i [\Lambda^k T^k]^{-1}$, is dependent on the *i*th LSF vector weights and KLT matrix and eigen values of the *k*th cluster. Now, if it is forced to assume that W^i is an identity matrix, then the weighting matrix can be simplified to, $O^{k,i} =$

 $\left[\left[\Lambda^{k}\right]^{-1}\right]^{T}\left[\Lambda^{k}\right]^{-1} = diag[\{\lambda_{j}^{k}\}_{j=1}^{16}]$ (as T^{k} is orthogonal). Therefore the distortion measure, characterized by Eqn. 3, is further simplified as:

$$d(\mathbf{x}^i, \hat{\mathbf{x}}^i) = \sum_{j=1}^{10} \lambda_j^k (u_j^i - \hat{u}_j^i)^2 \simeq d(\mathbf{u}^i, \hat{\mathbf{u}}^i)$$
(4)

where, $\{\lambda_j^k\}_{j=1}^{16}$ are the new cluster specific weighting coefficients. This modified WED measure is used at second stage to quantize the transformed coefficients of the *i*th frame LSF vector belonging to the *k*th cluster. Therefore, index of the cluster mean is found using ED measure at Q_1 , whereas the residual transformed vector, \mathbf{u}^i , is coded using the cluster specific WED measure given in Eqn 4 at Q_2 .

2.3. TSTVQ codebook training

The LBG algorithm is first applied on the full training database to produce M centroids (or mean vectors), $\{\mu^k\}_{k=1}^M$, using ED measure, which are the optimum code-vectors of Q_1 . All the training vectors are then classified based on the nearest neighbor criterion using ED measure and after that the cluster specific KLT matrices, $\{T^k\}_{k=1}^M$, are found out using classified data. Then, the training vectors are transformed using Eqn. 1 and pulled together to create a new training database to design SVQ codebook for the second stage quantizer Q_2 . The KLT is so ordered that the eigen values are in a descending order; hence the transformed vector, \mathbf{u}^i , is split into six parts as (2,2,2,3,3,4) sub-vectors in the second stage and quantized using SVQ technique with variance based bit allocation.

2.4. Computation and memory complexity of TSTVQ

The computational steps associated with TSTVQ method are: cluster search using ED measure at first stage, mean subtraction, KLT transformation, division by standard deviation values, finding cluster specific weights $(\{\lambda_{j}^{k}\}_{j=1}^{l-1})$ from standard deviation values, SVQ codebook search using WED measure at second stage, multiplication by standard deviation values, inverse KLT transformation and mean addition for reproduction. Using ED and WED measures, the codebook search complexity of a VQ, with *h*-dimensional vector and *B*-bit allocation, respectively are $3h2^{B} + 2^{B}$ flops¹ and $4h2^{B} + 2^{B}$ flops.

Let the dimension of the LSF vector is p (here, p = 16) and the bits allocated to Q_1 is b_0 (i.e. $M = 2^{b_0}$); the sub-vector dimensions at Q_2 are $\{p_i\}_{i=1}^6$ with corresponding bit allocations as $\{b_i\}_{i=1}^6$ such that $p = \sum_{i=1}^6 p_i$ and if the total bit allocation is b, then $b = \sum_{i=0}^6 b_i$. Therefore, the total required flops is: $(3p2^{b_0} + 2^{b_0}) + p + 2p^2 + p + p + (\sum_{i=1}^6 4p_i2^{b_i} + \sum_{i=1}^6 2^{b_i}) + p + 2p^2 + p$. On the other hand, for a five split traditional SVQ method (as implemented in [14]), if the sub-vector dimensions are $\{q_i\}_{i=1}^5$ with corresponding bit allocations as $\{c_i\}_{i=1}^5$ (such that $b = \sum_{i=1}^5 c_i$), then the necessary flops is: $\sum_{i=1}^5 4q_i2^{c_i} + \sum_{i=1}^5 2^{c_i}$. Now, we compare the memory requirements in terms of floats. For TSTVQ, the required memory to store the mean vectors, standard deviation values and KLT matrices of M clusters is: $2pM + p^2M$ floats. Also the required number of floats for SVQ codebook (in Q_2) storage is: $\sum_{i=1}^6 p_i 2^{b_i}$. On the other hand, traditional SVQ method needs $\sum_{i=1}^5 q_i 2^{c_i}$ floats for only codebook storage.

¹It is assumed that any operation, like summation, subtraction, multiplication, division or comparison, needs 1 floating point operation (flop).



Figure 2: Cluster index (value of 'k') of each frame to show interframe cluster continuity (ICC) property for consecutive 50 frames respectively using M = 16 and 32.

2.5. Inter-frame cluster continuity (ICC)

The VO/SVO and the KLT remove intra-frame redundancy of the LSF vector. However, the LSF parameters show a significant interframe correlation between successive frames since speech spectra are slowly time-varying. Therefore, it is observed that several consecutive frames' LSF vectors are in the same cluster and this property can be exploited to improve coding efficiency. This "interframe cluster continuity" (ICC) property is illustrated in Fig. 2 for 50 frames respectively using M = 16 and 32. In this method, a single bit is used to transmit the information that whether the cluster mean vector index of the current frame is same as the cluster mean vector index of the previous frame. Now, let b bits are allocated per frame. If the current frame LSF vector is in the same cluster of the previous frame, there is no need to transmit the cluster mean vector index and thus more bits can be spend to code \mathbf{u}^i ; therefore, the available bits to code \mathbf{u}^i is b-1. On the other hand when the current LSF vector is coded using a new cluster mean vector, the associated index must be transmitted and thus the available bits to code \mathbf{u}^i is $b - (1 + b_0)$. Therefore, it is necessary to store two SVQ codebooks for two different bit allocations at second stage and thus also the searching complexity at Q_2 differs from frame to frame depending on which of the two SVQ codebooks is used. Experimental results show that exploitation of ICC, improves the coding gain but at the expense of moderate increment of average computational complexity and memory requirement. We refer to the TSTVQ method along with ICC, as TSTVQICC method.

3. Quantization results

The TIMIT database is used in training and testing of the TSTVQ method, where speech is sampled at 16 kHz. We have used the specification of AMR-WB speech codec [12] to produce 16-th order LP coefficients which are then converted to LSF parameters. In the experiments, 368815 vectors are used for training and "out of training" 5000 frames are used for testing.

To measure the LSF quantization performance, we use the common measure of Spectral Distortion (SD) [3] and a recently proposed measure of Spectral Distortion with Interframe Memory (SDM) [10]. A low *average* SD and *rms* SDM along with minimum number of high SD outliers are necessary for good spectrum quantization performance ([3], [10]).

We investigate the quantization performance of TSTVQ method over direct two stage VQ (i.e. without KLT and variance normalization) and traditional SVQ methods. Table 1 shows the rate-distortion performance, computational complexity and memory requirements of the TSTVQ method at varying bit-rates and varying number of clusters at Q_1 . The bit allocation for Q_1 and



Table 1: Performance of TSTVQ at different number of clusters

| Tuble 1. 1 ent | Jinnane | 0 01 1 | 51 y Q m | uniteren | n number | or cruster | | | |
|--|-------------|--------|----------|----------|-----------|------------|--|--|--|
| Total bits/fi | rame | Avg. | >2 dB S | SD RM | S kflops/ | / floats/ | | | |
| {Bit allocati | on as: | SD | Outlier | s SDI | M frame | frame | | | |
| $(b_0, b_1, b_2 \dots$ | $. b_6) \}$ | (dB) | (in %) | (dB |) | (ROM) | | | |
| Number of clusters $M = 2^{b_0} = 16$ | | | | | | | | | |
| 42 (4,7,6,6,7 | 7,6,6) | 1.16 | 3.94 | 1.1 | 8 7.77 | 5952 | | | |
| 43 (4,7,6,6,8 | 8,6,6) | 1.12 | 3.14 | 1.14 | 4 9.44 | 6336 | | | |
| 44 (4,7,6,6,8 | 8,7,6) | 1.07 | 2.56 | 1.1 | 0 10.27 | 6528 | | | |
| 45 (4,7,6,6,8 | 8,7,7) | 1.03 | 2.10 | 1.0 | 6 11.36 | 6784 | | | |
| 46 (4,7,7,6,8 | 8,7,7) | 0.99 | 1.62 | 1.0 | 1 11.93 | 6912 | | | |
| Number of clusters $M = 2^{b_0} = 32$ | | | | | | | | | |
| 42 (5,6,6,6,7 | 7,6,6) | 1.14 | 3.82 | 1.1 | 6 7.95 | 10432 | | | |
| 43 (5,7,6,6,7 | 7,6,6) | 1.09 | 3.06 | 1.12 | 2 8.54 | 10560 | | | |
| 44 (5,7,6,6,7 | 7,7,6) | 1.05 | 2.62 | 1.0 | 8 9.37 | 10752 | | | |
| 45 (5,7,6,6,8 | 8,7,6) | 1.01 | 1.88 | 1.04 | 4 11.04 | 11136 | | | |
| 46 (5,7,6,6,8 | 8,7,7) | 0.98 | 1.62 | 1.0 | 1 12.12 | 11392 | | | |
| Number of clusters $M = 2^{b_0} = 64$ | | | | | | | | | |
| 42 (6,6,6,5,7 | 7,6,6) | 1.12 | 3.38 | 1.1 | 5 8.83 | 19548 | | | |
| 43 (6,6,6,6,7,6,6) | | 1.08 | 2.74 | 1.1 | 1 9.08 | 19648 | | | |
| 44 (6,6,6,6,8,6,6) | | 1.04 | 2.18 | 1.0 | 7 10.62 | 20032 | | | |
| 45 (6,6,6,6,8,7,6) | | 1.01 | 1.96 | 1.04 | 4 11.39 | 20224 | | | |
| 46 (6,7,6,6,8 | 8,7,6) | 0.97 | 1.64 | 1.0 | 0 11.90 | 20352 | | | |
| | | | | | | | | | |
| Table 2: Performance of TSTVQICC at $M = 2^{b_0} = 32$ | | | | | | | | | |
| Total bits/ | Avg. | >2 | dB SD | RMS | Avg. | floats/ | | | |
| frame | SD | Ou | tliers | SDM | kflops/ | frame | | | |
| | (dB) | (iı | n %) | (dB) | frame | (ROM) | | | |
| 42 | 1.09 | 3 | .16 | 1.12 | 9.64 | 12480 | | | |
| 43 | 1.05 | 2 | .60 | 1.08 | 10.19 | 12736 | | | |
| 44 | 1.01 | 2 | | 1.05 | 11.53 | 13376 | | | |
| 45 | 0.97 | 1 | .96 | 1.01 | 12.75 | 13952 | | | |

to the six sub-vectors of transformed coefficients at Q_2 , are also given. We observe that by increasing the number of clusters, lower distortion is achieved at all the bit-rates examined. The performance of the TSTVQICC method with different number of clusters, is also evaluated; Table 2 shows the results for M = 32.

0.97

13.90

14464

1.40

46

0.93

In case of traditional SVQ method, the 16 dimensional LSF vector is split into 5 parts of (3-3-3-3-4) sub-vectors [14] and the performance is shown in Table 3 with the bit allocations to sub-vectors. In this case we use the WED measure shown in Eqn. 2 where the weights are the popular inverse harmonic mean weights [8]. For the direct two stage VQ method the second stage residual vector is split into five parts of (3,3,3,3,4) sub-vectors and the performance of a two stage VQ method, with M = 32 (i.e. $b_0 = 5$ at Q_1) and with optimum bit allocations ($\{d_i\}_{i=1}^5$) to the five sub-vectors at Q_2 , is shown in Table 4 using appropriate distance measures. It may be noted that in all the methods, the codebooks are designed using the well-known LBG algorithm and ED measure.

Rate-distortion (average SD) performance and computational complexity of different methods is shown in Fig. 3 where the number of clusters, M = 32, is used for TSTVQ, TSTVQICC and two stage VQ methods. We observe that two stage VQ performs better than traditional SVQ method. The new TSTVQ performs better than two stage VQ and saves more than 2 bits compared to SVQ method even at much lower computational complexity, but at the expense of higher memory. Inclusion of cluster specific KLTs in

Table 3: Performance of traditional SVO method

| Total bits/frame | Avg. | >2 dB SD | RMS | kflops/ | floats/ |
|--------------------------|------|-----------|------|---------|---------|
| {Bit allocation as: | SD | Outliers | SDM | frame | frame |
| $(c_1, c_2 \dots c_5)\}$ | (dB) | (in %) | (dB) | | (ROM) |
| 42 (8,9,9,8,8) | 1.25 | 2.74 | 1.23 | 24.32 | 5632 |
| 43 (8,9,9,9,8) | 1.21 | 2.02 | 1.19 | 27.64 | 6400 |
| 44 (9,9,9,9,8) | 1.17 | 1.86 | 1.16 | 30.97 | 7168 |
| 45 (9,9,9,9,9) | 1.10 | 1.10 | 1.10 | 35.32 | 8192 |
| 46 (9,10,9,9,9) | 1.06 | 0.90 | 1.06 | 41.98 | 9728 |

Table 4: Performance of two stage VQ method at $M = 2^{b_0} = 32$

| Total bits/frame | Avg. | >2 dB SD | RMS | kflops/ | floats/ |
|-------------------------------|------|-----------|------|---------|---------|
| {Bit allocation as: | SD | Outliers | SDM | frame | frame |
| $(b_0, d_1, d_2 \dots d_5)\}$ | (dB) | (in %) | (dB) | | (ROM) |
| 42 (5,7,8,8,7,7) | 1.19 | 3.08 | 1.20 | 13.76 | 3328 |
| 43 (5,7,8,8,8,7) | 1.15 | 2.26 | 1.16 | 15.42 | 3712 |
| 44 (5,8,8,8,8,7) | 1.12 | 2.10 | 1.13 | 17.08 | 4096 |
| 45 (5,8,8,8,8,8) | 1.05 | 1.12 | 1.07 | 19.26 | 4608 |
| 46 (5,8,9,8,8,8) | 1.02 | 0.94 | 1.03 | 22.59 | 5376 |

TSTVQ method mitigates the coding loss due to splitting of residual vector and thus allows independent VQs of smaller dimensions in the second stage, resulting in better performance and lower computational complexity compared to the two stage VQ and SVQ methods. Exploitation of ICC along-with TSTVQ, in TSTVQICC, further improves the coding gain and saves 3-4 bits compared to the traditional SVQ method; TSTVQICC also reduces the computational complexity by 58-66% compared to the SVQ method, but needs 1.5-2.2 times of memory.

4. Conclusions

We explore the use of a two stage transform VQ method, for wideband speech LSF quantization, where cluster specific KLT and variance normalization are used to design an efficient common SVQ codebook in the second stage. The advantage of higher dimensional coding efficiency of VQ is exploited at the first stage and the coding loss of splitting the residual vector is mitigated using a decorrelating transform [11] at the second stage. A cluster specific WED measure is derived for efficient VQ encoding in the second stage. Further coding gain is achieved by exploiting inter-frame cluster continuity property. It is shown that proposed TSTVQICC method performs better than two stage VQ method and saves 3-4 bits than traditional SVQ method even at a much lower computational complexity.

5. References

- N. Phamdo and N. Farvardin, "Coding of speech LSP parameters using TSVQ with interblock noiseless coding", Proc. ICASSP, vol.1, pp. 193-196, April, 1990.
- [2] A. Gersho and R. M. Gray, "Vector Quantization and Signal Compression," Kluwer Academic, Ch. 12, 1992.
- [3] K. K. Palliwal and B. S. Atal, "Efficient vector quantization of LPC parameters at 24 bits/frame", IEEE Trans. Acoust., Speech, Signal Proc., vol. 1, pp. 3-14, Jan 1993.
- [4] W. F. LeBlanc, B. Bhattacharya, S. A. Mahmoud and V. Cuperman, "Efficient search and design procedures for robust multi-stage VQ of LPC parameters for 4 Kb/s speech cod-



Figure 3: Comparison of rate-distortion (average SD) performance with computational complexity for different methods. TSTVQICC provides 3-4 bits advantage and reduces the computational complexity by 58-66% compared to SVQ.

ing", IEEE Trans. Speech and Audio Proc., Vol. 1, No. 4, pp. 373-385, October 1993.

- [5] A. Ubale and A.Gersho, "A multi-band CELP wideband speech coder", Proc. ICASSP, pp. 1367-1370, 1994.
- [6] R. Lefebvre, R. Salami, C. Laflamme and J. P. Adoul, "High quality coding of wideband audio signals using transform coded excitation (TCX)", Proc. ICASSP, pp. 193-196, 1994.
- [7] J. H. Chen and D. Wang, "Transform predictive coding of wideband speech signals", Proc. ICASSP, pp. 275-278, 1996.
- [8] H. L. Vu and L. Lois, "Efficient distance measure for quantization of LSF and its Karhunen-Loeve transformed parameters", IEEE Trans. Speech and Audio Proc., vol. 8, No.6, pp. 744-746, Nov 2000.
- [9] G. Guibe, H. T. How and L. Hanzo, "Speech spectral quantizers for wideband speech coding", European Trans. on Telecommunications, 12(6), pp. 535-545, 2001.
- [10] F. Norden and T. Eriksson, "Time evolution in LPC spectrum coding", IEEE Trans. Speech and Audio Proc., vol.12, No.3, pp. 290-301, May 2004.
- [11] F. Norden and T. Eriksson, "On split quantization of LSF parameters", Proc. ICASSP, vol.1, pp. I-157-160, May, 2004.
- [12] "AMR wide-band speech codec, transcoding functions (Release 5):" 3GPP TS 26.190 V 5.1.0
- [13] S. So and K.K.Paliwal, "Multi-frame GMM-based block quantisation of line spectral frequencies for wide-band speech coding", Proc. ICASSP, vol.1, pp. 121-124, March, 2005.
- [14] S. So and K.K.Paliwal, "Switched split vector quantisation of line spectral frequencies for wide-band speech coding", Proc. Eurospeech, Interspeech-2005, Lisbon, Portugal, pp. 2705-2708, Sept. 2005.