

Quantização vetorial particionada com renormalização

Miguel Arjona Ramírez

Resumo—Propõe-se uma variante do algoritmo de quantização vetorial particionada (QVP) que reduz a perda por particionamento, mantendo a baixa complexidade de busca que caracteriza a QVP. Ele atua em parte nas fases de treinamento e de codificação, quando normaliza a largura de faixa que cada subvetor ocupa. Por outra parte, na decodificação cada subvetor selecionado do dicionário de códigos tem sua largura de faixa renormalizada para cobrir ao máximo a região permitida pela restrição de ausência de sobreposição entre faixas de subvetores vizinhos. Considerando esses procedimentos, o algoritmo é denominado de QVP com renormalização (QVPR). O desempenho da QVPR é analisado em comparação com a QVP comum na codificação de vetores de frequências das raias espectrais (LSFs) que representam a envoltória espectral de curto prazo de sinais de voz de faixa larga, para os quais a QVPR economiza quatro bits por vetor, atingindo codificação transparente em 42 bit/quadro.

Palavras-Chave—Quantização vetorial particionada (QVP), frequências das raias espectrais (LSFs), análise de sinais de voz, codificação de voz.

Abstract—A variant split vector quantization (SVQ) algorithm is proposed that reduces the split loss while keeping the complexity low as usual for a split algorithm. Partially, it operates in the training and encoding phases, when it adjusts the width of the frequency band spanned by each split. Complementarily, in the decoding phase each subvector selected from a split of the codebook undergoes a bandwidth renormalization in order to maximally cover the region allowed by a nonoverlapping constraint between neighboring split bands. In consideration to these procedures, the algorithm is called renormalized SVQ (RSVQ). The performance of RSVQ is analyzed in comparison to standard SVQ in coding line spectral frequency (LSF) vectors that represent the short-term spectral envelope of wideband speech, where RSVQ saves four bits per vector, reaching transparent coding at 42 bit/frame.

Keywords—Split vector quantization (SVQ), line spectral frequencies (LSFs), speech analysis, speech coding.

I. INTRODUÇÃO

A quantização vetorial é mais eficiente que a quantização escalar em termos de taxa e distorção. Entretanto, sua correspondente complexidade de busca cresce exponencialmente com a dimensão do vetor para a busca exaustiva. Por sorte, há métodos subótimos de alta eficiência como a quantização vetorial particionada [1], que é comumente usada para a codificação da envoltória espectral de sinais de voz.

A vantagem da QVP é particularmente grande quando a envoltória espectral de curto prazo é representada por co-

eficientes LSF porque, quando a perturbação atinge apenas um deles, a distorção aparece na região espectral localizada em torno do seu valor [2]. Este comportamento está de acordo com a percepção auditiva humana e é razoavelmente operacionalizado pela distorção espectral (SD) logarítmica, como explicado posteriormente na Seção II.

Por outro lado, a partição dos vetores impossibilita a detecção de correlações entre subvetores. Este efeito é chamado de “perda por particionamento” [3].

Várias combinações de métodos foram propostas para melhorar o desempenho da QV além da QVP como a quantização vetorial particionada multiestágio (S-MSVQ), usada no codificador de voz AMR-WB [4], ou a quantização vetorial particionada chaveada (SSVQ) [5]. Aquele algoritmo aplica QVP e QV multiestágio ao sinal residual da predição de média móvel da sequência de vetores LSF, enquanto este algoritmo usa codificação por transformadas e QV classificada.

Mesmo sendo eficientes estes métodos, a S-MSVQ pode prolongar os efeitos de erros no canal através de seu preditor. Ao contrário, na QVPR um processo intraquadro reduz a perda por particionamento através da normalização da largura de faixa de cada subvetor antes de sua quantização na fase de treinamento e no processo de codificação de um vetor teste. Este processo guarda semelhanças com a normalização descrita em [6] para voz de faixa telefônica estreita, que também é usado em [7] conjuntamente com QV classificada. Então, no ato da decodificação de um vetor teste de LSFs, a largura de faixa de cada subvetor é renormalizada, como relatado na Seção III. Antes disso, são introduzidas medidas de distorção para o projeto e avaliação de quantizadores vetoriais e relacionadas com particionamento da largura de faixa na Seção II. Depois disso, são apresentados resultados experimentais, também discutidos na Seção IV antes da conclusão.

II. MEDIDAS DE DISTRORÇÃO E MODOS DE PARTIÇÃO DA FAIXA DE FREQUÊNCIA

As medidas espectrais podem ser razoavelmente adaptadas para realçar aspectos auditivos perceptivos para o projeto e avaliação de codificadores. De fato, dado o vetor de LSFs $\mathbf{f} = [f_1 \ f_2 \ \dots \ f_p]^T$, em que p é a ordem de predição, e sua versão quantizada $\tilde{\mathbf{f}} = [\tilde{f}_1 \ \tilde{f}_2 \ \dots \ \tilde{f}_p]^T$, a medida que é mais empregada para avaliação de quantizadores é a distorção espectral (SD) logarítmica entre eles, definida como

$$D_S(\mathbf{f}, \tilde{\mathbf{f}}) = \sqrt{\int_0^1 (10 \log_{10} P(f) - 10 \log_{10} \tilde{P}(f))^2 df}, \quad (1)$$

Miguel Arjona Ramírez, Departamento de Engenharia de Sistemas Eletrônicos, Escola Politécnica, Universidade de São Paulo, São Paulo, Brasil, E-mail: miguel@lps.usp.br. Este trabalho é financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) através do Processo no. 309249/2008-2 e pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) através do Processo no. 2009/18242-5.

sendo $P(f)$ e $\tilde{P}(f)$ as densidades espectrais de potência (DEPs) dos filtros de síntese associados aos vetores LSF \mathbf{f} e $\tilde{\mathbf{f}}$, respectivamente, e f é a frequência cíclica em ciclo/amostra.

A maior razão para a popularidade da SD está relacionada com sua aplicação nas três condições seguintes para quantização vetorial de informação sobre os coeficientes de predição linear (LPC), enunciadas por Paliwal e Atal [2]:

- 1) O valor médio da SD aproxima-se de 1 dB,
- 2) Não há nenhum quadro com distorção excedente que tenha SD maior que 4 dB,
- 3) O número de quadros com distorção excedente com SD na faixa 2–4 dB é menor que 2%.

Na QVP o vetor de LSFs é particionado como

$$\mathbf{f} = [\varphi_1^T \quad \varphi_2^T \quad \cdots \quad \varphi_\zeta^T]^T,$$

sendo ζ o número de partes e o i -ésimo subvetor consiste em

$$\varphi_i = [f_{\delta_i} \quad f_{\delta_i+1} \quad \cdots \quad f_{\delta_i+D_i-1}]^T \quad (2)$$

com índice de LSF inicial δ_i e dimensão D_i com índices de LSF extremos $\delta_1 = 1$ e $\delta_\zeta + D_\zeta - 1 = p$.

Entretanto, para projetar o quantizador vetorial, a medida SD é costumeiramente complexa demais para ser usada. Em geral, usa-se a distância quadrática ponderada no seu lugar. Em particular, usamos a seguinte medida quadrática dinamicamente ponderada [2]

$$d(\mathbf{f}, \tilde{\mathbf{f}}) = (\mathbf{f} - \tilde{\mathbf{f}})^T \mathbf{W} (\mathbf{f} - \tilde{\mathbf{f}}) \quad (3)$$

sendo \mathbf{W} a matriz de ponderação diagonal com elementos na diagonal principal

$$w_{ii} = P^r(f_i) \quad (4)$$

para $i = 1, 2, \dots, p$, em que o expoente da DEP é fixado em $r = 0, 3$.

É importante notar que a medida de distorção na Eq. (3) é separável em subvetores componentes, ou seja,

$$d(\mathbf{f}, \tilde{\mathbf{f}}) = \sum_{i=1}^{\zeta} d_i(\varphi_i, \tilde{\varphi}_i), \quad (5)$$

sendo que a componente da distorção devida ao i -ésimo subvetor é

$$d_i(\varphi_i, \tilde{\varphi}_i) = \sum_{j=\delta_i}^{\delta_i+D_i-1} w_{jj} (f_j - \tilde{f}_j)^2. \quad (6)$$

III. NORMALIZAÇÃO E RENORMALIZAÇÃO DAS LARGURAS DE FAIXA DOS SUBVETORES

A separabilidade da medida de distorção dinamicamente ponderada em partes implica que o projeto de um quantizador vetorial particionado pode ser contido dentro de cada classe de subvetores. Em outras palavras, não é necessário que a distorção acumulada seja transferida do projeto do quantizador de uma classe de subvetores para outra.

De outro ponto de vista, não é útil no projeto do quantizador para uma classe de subvetores o emprego da distorção acumulada ao longo dos projetos para as classes de subvetores de frequências inferiores. Impõem-se, então, a indagação sobre

a forma como a informação entre as classes de subvetores poderia ser comunicada entre os projetos.

Uma resposta surge quando se percebe que uma grande parte da perda por particionamento se origina da redução de tamanho do subdicionário imposta pela aplicação das relações de estabilidade

$$0 < \tilde{f}_1 < \tilde{f}_2 < \cdots < \tilde{f}_{p-1} < \tilde{f}_p < 1/2, \quad (7)$$

que excluem de comparação os subvetores-código que não as satisfazem. Por exemplo, na codificação do i -ésimo subvetor, apenas são comparados os subvetores que satisfazem

$$f_{\delta_i} > \tilde{f}_{\delta_i-1}, \quad (8)$$

isto é, a mais baixa LSF no subvetor tem que ser maior que a mais alta LSF quantizada no subvetor anterior. A proporção de exclusão pode ser maior que um quinto do tamanho do dicionário de códigos conforme analisado na Seção IV a partir de experimentos.

Numa possibilidade de normalização do i -ésimo subvetor, sua frequência extrema inferior assume o valor quantizado da maior LSF do subvetor imediatamente inferior do quadro m atual na base de dados de treino, ou do quadro a ser codificado, como

$$f_{iL}(m) = \tilde{f}_{\delta_i-1}(m), \quad (9)$$

e seja $f_{iU} = 1/2$ a frequência extrema superior. Denominaremos esta versão da QVPR de QVPR sequencial. Então o i -ésimo subvetor LSF a ser quantizado é normalizado como

$$f'_j(m) = \frac{f_j(m) - f_{iL}(m)}{f_{iU} - f_{iL}(m)} \quad (10)$$

para $j = \delta_i, \delta_i + 1, \dots, \delta_i + D_i - 1$. Em seguida, o subvetor normalizado \mathbf{f}' é quantizado como $\tilde{\mathbf{f}}'$, selecionado do i -ésimo subdicionário de códigos.

Na fase de decodificação, a frequência extrema inferior para renormalização do i -ésimo subvetor é idêntica à LSF quantizada mais alta do subvetor anterior do mesmo vetor $\tilde{\mathbf{f}}'(m)$, sendo

$$\tilde{f}_{iL}(m) = \tilde{f}_{\delta_i-1}(m) \quad (11)$$

para ser usada na renormalização que gera o subvetor quantizado $\tilde{\mathbf{f}}$ como

$$\tilde{f}_j(m) = (f_{iU} - \tilde{f}_{iL}(m)) \tilde{f}'_j(m) + \tilde{f}_{iL}(m) \quad (12)$$

para $j = \delta_i, \delta_i + 1, \dots, \delta_i + D_i - 1$.

Outra possibilidade de normalização é a QVPR entrelaçada, que usa os dois subvetores adjacentes ao que está em quantização para determinar a frequência extrema inferior do i -ésimo subvetor no m -ésimo quadro pela Eq. (9), de forma semelhante ao caso sequencial, e a frequência extrema superior é determinada por

$$f_{iU}(m) = \tilde{f}_{\delta_{i+1}}(m), \quad (13)$$

isto é, a frequência extrema superior é a LSF mais baixa do subvetor seguinte.

De forma semelhante, a renormalização na fase de decodificação usa a mesma atribuição de valor para a frequência extrema inferior do i -ésimo subvetor no m -ésimo

quadro como descrito pela Eq. (11), mas, ao contrário, a frequência extrema superior é adaptativa e ajustada por

$$\tilde{f}_{iU}(m) = \tilde{f}_{\delta_i+1}(m) \quad (14)$$

para ser usada na normalização do vetor-código selecionado $\tilde{\varphi}'_i(m)$ de acordo com

$$\tilde{f}_j(m) = (\tilde{f}_{iU}(m) - \tilde{f}_{iL}(m)) \tilde{f}'_j(m) + f_{iL}(m) \quad (15)$$

para $j = \delta_i, \delta_i + 1, \dots, \delta_i + D_i - 1$. Naturalmente, tanto o $(i-1)$ -ésimo como o $(i+1)$ -ésimo subvetor precisam ter sido quantizados antes pela QVP comum.

Num experimento relatado na Seção IV, a versão sequencial mostrou-se superior à versão entrelaçada por uma pequena margem de diferença.

IV. ANÁLISE EXPERIMENTAL

Os experimentos mais importantes foram destinados à comparação do algoritmo QVPR com o algoritmo QVP comum na quantização de vetores de coeficientes LSF de voz de faixa larga. Como foram apresentadas duas versões de sequências de partição, a sequencial e a entrelaçada, na Seção III, um experimento preliminar compara os desempenhos das duas.

Os sinais de voz foram tomados da base TIMIT de sinais de voz [8], [9] e analisados pelo codificador de voz de faixa larga AMR-WB do consórcio 3GPP [10], [4] e recomendado pela ITU-T como o codificador de voz G.722.2 [11]. Desta forma, foram gerados vetores LSF de dimensão 16 a uma taxa de 50 quadro/s com quadros de 30 ms de duração ponderados por uma janela de Hamming assimétrica. A partição de treino, com 705580 quadros, foi usada para o projeto de dicionários de QV enquanto a partição de teste, com 257852 quadros, foi usada para testar os quantizadores vetoriais.

Para os testes de todos os algoritmos de QV particionada, os vetores de LSFs de dimensão 16 foram partidos em cinco subvetores $\{\varphi_i\}_{i=1}^5$ de dimensões 3,3,3,3,4, respectivamente. A perda por particionamento na QVP foi avaliada em termos da disponibilidade de dicionário de códigos resultante da aplicação das restrições de estabilidade a cada subdicionário de códigos na quantização dos subvetores de teste em sete taxas de quadro compostas, mostradas na Tabela I. Desconsiderando o primeiro subvetor, que trivialmente satisfaz as relações de estabilidade, pelo menos 14% dos vetores-código e não mais que 22% dos vetores-código podem resultar inúteis durante a quantização.

De um ponto de vista diferente, as superposições entre as faixas adjacentes a partir das frequências dos pontos de partição podem ser visualizadas pelas funções de densidade de probabilidade marginais ao longo da partição de treino, visualizadas em pares de frequências extremas de subvetores adjacentes na Fig. 1. As sobreposições mais extensas são observadas entre ambas extremidades de uma mesma classe de subvetores e entre a frequência extrema superior de uma classe de subvetores e a frequência extrema inferior da classe de subvetores seguinte, mas mesmo a frequência extrema inferior de uma classe de subvetores e a frequência extrema superior da classe de subvetores seguinte podem se tocar.

TABELA I
DISPONIBILIDADE DE DICIONÁRIO DE CÓDIGOS RESULTANTE DA APLICAÇÃO DAS RESTRIÇÕES DE ESTABILIDADE À QVP COMUM PARA VETORES LSF DE DIMENSÃO 16 DIVIDIDOS EM SUBVETORES DE DIMENSÕES (3,3,3,3,4).

Taxa		Disponibilidade de dicionário por subdicionário (%)				
Por quadro (bit/fr)	Por subvetor (bit/sp)					
40	(8,8,8,8,8)	100	85,5	77,6	80,4	79,9
41	(8,9,8,8,8)	100	85,5	77,6	80,4	79,9
42	(8,9,9,8,8)	100	85,6	77,6	80,4	79,9
43	(8,9,9,9,8)	100	85,6	77,6	80,4	79,9
44	(9,9,9,9,8)	100	85,5	77,6	80,4	79,9
45	(9,9,9,9,9)	100	85,5	77,6	80,4	79,9
46	(9,10,9,9,9)	100	85,5	77,5	80,4	79,8

Em primeiro lugar, fez-se um experimento para comparar o efeito da renormalização sobre a QVP e também para avaliar as versões sequencial e entrelaçada da QVPR. Foi executado à taxa de 46 bit/quadro, particionada como 9,10,9,9,9 bit/subvetor para os subvetores de 1 a 5, respectivamente. Esta taxa foi selecionada porque se havia verificado que a QVP comum permite quantização transparente nesta taxa.

O algoritmo QVPR quantiza sequencialmente os subvetores de 1 a 5 nesta ordem ao passo que a versão entrelaçada inicia quantizando os subvetores intermediários 2 e 4 através da QVP comum para então prosseguir com a quantização dos subvetores 1, 3 e 5 pela QVPR propriamente dita.

Os resultados são apresentados na Tabela II, em que se pode observar que os algoritmos com renormalização são significativamente superiores ao algoritmo comum, tanto em SD média como em número de excedentes no intervalo 2–4 dB. Além disso, o algoritmo QVPR sequencial tem um desempenho ligeiramente superior ao da QVPR entrelaçada, tendo sido escolhido para a sequência de testes mais extensivos seguinte.

TABELA II
DESEMPENHOS DAS QVPRS SEQUENCIAL E ENTELÇADA E DA QVP COMUM PARA VETORES LSF DE DIMENSÃO 16 PARTIDOS EM SUBVETORES DE DIMENSÃO (3,3,3,3,4) NUMA TAXA DE 46 BIT/QUADRO, INCLUINDO A DISTORÇÃO ESPECTRAL LOGARÍTMICA E DUAS CLASSES DE EXCEDENTES.

Algoritmo	Média SD (dB)	Excedentes	
		2 – 4 dB (%)	> 4 dB (ppm)
QVP comum	0,924	0,46	0
QVPR sequencial	0,821	0,22	0
QVPR entrelaçada	0,837	0,22	0

A QVP comum e a QVPR proposta foram testadas numa sequência de taxas de quadro de 40 bit/quadro a 46 bit/quadro, obtendo-se os resultados apresentados nas Tabelas III e IV, respectivamente. Pode-se notar que o algoritmo QVP perde sua transparência à medida que a taxa cai abaixo de 46 bit/quadro devido a sua distorção espectral gradualmente crescente. Mais precisamente, de acordo com as condições para quantização transparente, a causa principal para a perda da transparência em taxa imediatamente abaixo de 46 bit/quadro é o número não-nulo de quadros excedendo 4 dB de distorção.

Por outro lado, o algoritmo QVPR cobre toda a faixa de

TABELA III

DESEMPENHO DA QVP COMUM PARA VETORES LSF DE DIMENSÃO 16 EM PARTES DE DIMENSÕES (3,3,3,3,4), INCLUINDO DISTORÇÃO ESPECTRAL LOGARÍTMICA MÉDIA E DUAS CLASSES DE EXCEDENTES.

Taxa		Média SD (dB)	Excedentes	
Por quadro (bit/q)	Por subvetor (bit/sub)		2 – 4 dB (%)	> 4 dB (ppm)
40	(8,8,8,8,8)	1,175	2,55	19
41	(8,9,8,8,8)	1,127	1,92	19
42	(8,9,9,8,8)	1,072	1,15	40
43	(8,9,9,9,8)	1,046	1,09	40
44	(9,9,9,9,8)	1,019	1,01	40
45	(9,9,9,9,9)	0,962	0,59	40
46	(9,10,9,9,9)	0,924	0,46	0

TABELA IV

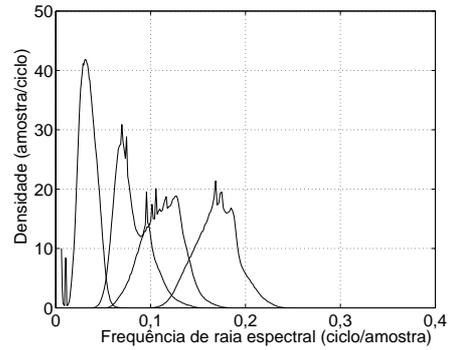
DESEMPENHO DA QVP COM RENORMALIZAÇÃO PARA VETORES LSF DE DIMENSÃO 16 EM PARTES DE DIMENSÕES (3,3,3,3,4), INCLUINDO DISTORÇÃO ESPECTRAL LOGARÍTMICA MÉDIA E DUAS CLASSES DE EXCEDENTES.

Taxa		Média SD (dB)	Excedentes	
Por quadro (bit/q)	Por subvetor (bit/sub)		2 – 4 dB (%)	> 4 dB (ppm)
40	(8,8,8,8,8)	1,047	1,29	0
41	(8,9,8,8,8)	1,003	0,96	0
42	(8,9,9,8,8)	0,968	0,75	0
43	(8,9,9,9,8)	0,938	0,59	0
44	(9,9,9,9,8)	0,907	0,51	0
45	(9,9,9,9,9)	0,855	0,26	0
46	(9,10,9,9,9)	0,821	0,22	0

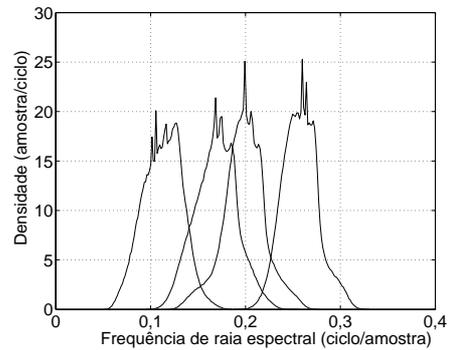
taxas de codificação de 46 bit/quadro até 40 bit/quadro sem nunca gerar um único quadro excedente de 4 dB de distorção. A taxa limiar da codificação transparente é determinada pela SD média na marca de 42 bit/quadro com boa margem de segurança. Em comparação com SSVQ e S-MSVQ [5], os resultados apresentados na Tabela IV são melhores tanto em SD média quanto em número de excedentes.

V. CONCLUSÃO

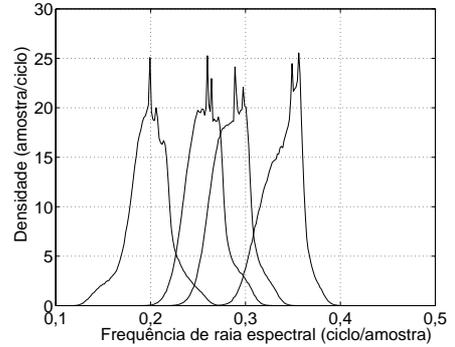
Uma variante da quantização vetorial particionada que opera em dois estágios foi proposta para reduzir a perda por particionamento e foram apresentados resultados de sua aplicação na quantização da envoltória espectral de curto prazo de sinais de voz de faixa larga representados pelas suas LSFs. O primeiro estágio ocorre no treinamento e na codificação, quando a faixa coberta através de cada subvetor é normalizada de forma tal que todos os vetores-códigos resultantes satisfazem as relações de ordenação que garantem a estabilidade do filtro de síntese. Posteriormente, o segundo estágio acontece no decodificador, quando os subvetores-código selecionados são renormalizados de forma tal que os subvetores que hão de compor o vetor teste quantizado se ajustem de acordo com as relações de ordenação. Os resultados encontrados são melhores que aqueles reportados para métodos mais complicados de QV classificada não-preditiva e para algoritmos preditivos, tanto em termos de valores médios da distorção espectral logarítmica quanto em número de excedentes.



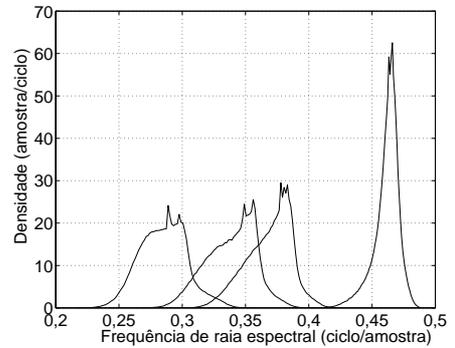
(a) Funções de densidade dos extremos dos subvetores 1 e 2.



(b) Funções de densidade dos extremos dos subvetores 2 e 3.



(c) Funções de densidade dos extremos dos subvetores 3 e 4.



(d) Funções de densidade dos extremos dos subvetores 4 e 5.

Fig. 1. Funções de densidade marginal para as frequências extremas dos subvetores adjacentes na partição de treino da base de dados.

REFERÊNCIAS

- [1] Allen Gersho and Robert M. Gray, *Vector quantization and signal compression*, Kluwer Academic Publishers, Boston, 1992.
- [2] Kuldip K. Paliwal and Bishnu S. Atal, "Efficient vector quantization of LPC parameters at 24 bits/frame," *IEEE Trans. Speech Audio Processing*, vol. 1, no. 1, pp. 3–14, Jan. 1993.
- [3] Saikat Chatterjee and Thippur V. Sreenivas, "Sequential split vector quantization of LSF parameters using conditional PDF," in *Proc. of IEEE Int. Conf. Acoust., Speech, Signal Processing*, Honolulu, 2007, vol. 4, pp. IV-1101–IV-1104.
- [4] 3rd Generation Partnership Project, "*Technical Specification Group Services and System Aspects; Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; ANSI-C code*," 3GPP TS 26.204 V7.1.0 (2007-09) Release 7, 2007.
- [5] Stephen So and Kuldip K. Paliwal, "A comparative study of LPC parameter representations and quantisation schemes for wideband speech coding," *Digital Signal Processing*, vol. 17, pp. 114–137, 2007.
- [6] Aaron M. Smith, James P. Ashley, Mark A. Jusiuk, and Weimin Peng, "Normalization and polygon error detection for split VQ of line spectral frequencies," in *Proc. of IEEE Workshop on Speech Coding*, Delavan, 2000, pp. 123–125.
- [7] Hosang Sung, Changyong Son, Kangeun Lee, and Eunmi Oh, "An efficient LSF quantization using dynamic bit allocation," in *Proc. of IEEE Int. Conf. Acoust., Speech, Signal Processing*, Taipei, Taiwan, 2009, pp. 4141–4144.
- [8] W. M. Fisher, G. R. Doddington, and K. M. Goudie-Marshall, "The DARPA speech recognition research database: specifications and status," in *Proc. DARPA Speech Recognition Workshop*, Feb. 1986, vol. 1, pp. 93–99.
- [9] John S. Garofolo, Lori F. Lamel, William M. Fisher, Jonathan G. Fiscus, David S. Pallett, Nancy L. Dahlgren, and Victor Zue, "TIMIT acoustic-phonetic continuous speech corpus," Linguistic Data Consortium, 1993, <http://www.ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogID=LDC93S1>.
- [10] Bruno Bessette, Redwan Salami, Roch Lefebvre, Milan Jelínek, Jani Rotola-Pukkila, Janne Vainio, Hannu Mikkola, and Kari Järvinen, "The adaptive multirate wideband speech codec (AMR-WB)," *IEEE Trans. Speech Audio Processing*, vol. 10, no. 8, pp. 620–636, Nov. 2002.
- [11] ITU-T, "*G.722.2: Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)*," Recommendation G.722.2, July 2003, <http://www.itu.int/rec/T-REC-G.722.2/en>.