A Smoothed and Thresholded Linear Prediction Analysis for Efficient Speech Coding

Aadel Alatwi and Kuldip K. Paliwal

Signal Processing Laboratory, Griffith University, Brisbane, QLD 4111, Australia

Email: aadel.alatwi@griffithuni.edu.au; k.paliwal@griffith.edu.au

Abstract —In this paper, we propose a new method of linear prediction (LP) analysis for estimating LP coefficients that are used in current speech coders. This method improves the robustness of LP coefficients computed from speech signals corrupted by noise as well as offering better quantisation efficiency. The quantisation performance and the noise-robustness obtained by the proposed LP coefficients were compared to that obtained by the LP coefficients computed using the autocorrelation and LP spectrum modification methods of LP Analysis, in terms of spectral distortion (SD). The results indicate that the proposed LP coefficients were more robust to noise and also offered transparent quantisation at lower bit-rates (savings of up to 2 bits/frame) than other LP coefficients.

Index Terms—Linear prediction coefficients, linear prediction analysis, two-split vector quantisation, three-split vector quantisation, spectral distortion measure

I. INTRODUCTION

The autocorrelation method of linear prediction analysis (AM-LP) is widely used as a basic technique for low bit-rate speech coding applications [1]. In these applications, obtained-LP coefficients, which represent the power spectrum envelope (short-term) of a speech signal using a low-order all-pole filter [2], are typically converted to LP parameters (e.g. line spectral frequencies (LSFs)) and then quantised using as few bits as possible prior to their transmission [3]. In noise-free environments, the performance of LP parameter-based speech coders is often satisfactory. However, in the presence of background noise, the LP analysis method leads to an inaccurate estimate of the LP spectrum of the speech; hence, the variance of the estimated LP coefficients is increased [4], which results in an overall quality deterioration of the reconstructed speech [5] and [6].

Various studies have been proposed to improve the LP analysis's immunity to noise. A noise reduction method using the technique of pitch synchronous addition was reported in [7]. In this approach, multiple pitch periods taken from the analysis frame are synchronously averaged to obtain the estimated frame. The analysis of pitch synchronous yields reliable estimation of pitch in a voiced segment of speech signal, however, it is susceptible to errors at the boundaries of unvoiced and voiced segments where the pitch periods are often irregular. The noise compensation approach, which is employed in the LP analysis of the noisy speech, was reported in [6]. The subtraction procedure of this method is applied iteratively to ensure that the autoregressive (AR) model is stable, in this case the Levinson-Durbin algorithm is used for each iteration. An LP spectrum modification approach (LPSM) based on linear extrapolation for noisy speech was proposed in [8]. In this scheme, the LP spectrum is modified in the high frequency region by applying a linear suppression rule, while the low frequency region remains preserved. The Spectral Envelope Estimation Vocoder (SEEVOC) method is another scheme proposed to improve the conventional LP analysis method's performance [9]. In the SEEVOC method, only those portions of the speech power spectrum obtained by the Fast Fourier Transform (FFT), which are less noisy, are used. Therefore, it attempts to clean the noisy speech power spectrum by ignoring the spectral parts that are affected more by noise. In terms of the estimation of the power spectrum envelope, the accuracy of this method requires a prior knowledge of the average speech signal pitch (for aperiodic waveforms). The behavior of the aforementioned methods is not appropriate for real-time applications.

In this paper, a modified procedure for performing LP analysis is proposed which reduces the variance of LP coefficient estimates in order to provide some noise robustness and also exploits some properties of the human auditory system. The LP coefficients are fully compatible with current speech coding standards and the algorithm proposed can be easily incorporated into the implementations of existing code. The quantisation performance of the proposed LP parameters was evaluated using different split vector quantisation schemes, as well as their robustness to noise in terms of the spectral distortion (SD) measurement. The results indicate that the proposed method provides a more accurate and noise-robust estimation of the LP parameters.

This paper is organized as follows: Section II describes the principles of the conventional LP analysis method. Section III provides the framework for the proposed LP analysis method. Section IV outlines the experimental setup that we used to measure the quantisation performance and the robustness of the LP parameters. Section V clarifies the details of each experiment and

provides the results. Section VI presents the conclusion of this study.

II. LP ANALYSIS OF SPEECH SIGNAL

In the theory of LP analysis [1], the speech signal $x(n)$ can be approximately predicted by a linear combination of the past p samples $\{x(n-1), x(n-2), x(n-3), \ldots, x(n-p)\}$, which can be described as:

$$\hat{x}(n) = -\sum_{k=1}^{p} a_k x(n-k) \tag{1}$$

where a_k is the LP coefficients and $\hat{x}(n)$ is referred to as a prediction of $x(n)$. Furthermore, a pth all-pole filter $H(z)$ can be used to represent the spectral envelope of speech signals $x(n)$, which is given by:

$$H(z) = \frac{G}{1 + \sum_{k=1}^{p} a_k z^{-k}} \tag{2}$$

where filter gain G is to conserve the total energy between the speech and impulse response of $H(z)$. The LP coefficients $\{a_k; k = 1, 2, \ldots, p\}$ and the gain are estimated by solving the Yule-Walker equations [10]:

$$\sum_{k=1}^{p} R(j-k)a_k = -R(j), \quad \text{for } j = 1, 2, \ldots, p \tag{3}$$

$$G = R(0) + \sum_{k=1}^{p} a_k R(k) \tag{4}$$

where $R(k)$ are the autocorrelation coefficients estimated from a frame of N samples of the speech signal $x(n)$ [10]:

$$R(k) = \frac{1}{N} \sum_{k=0}^{N-1-k} x(n)x(n+k) \tag{5}$$

Another method of estimating autocorrelation coefficients utilises the Einstein-Wiener-Khintchine theorem [10], by taking the inverse discrete-time Fourier transform (IFFT) of the periodogram estimate of the power spectrum $P(\omega)$:

$$R(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} P(\omega)e^{j\omega k} d\omega \tag{6}$$

where

$$P(\omega) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x(n)e^{-j\omega n} \right|^2 \tag{7}$$

This relationship between the periodogram and autocorrelation coefficients motivates our method for reducing the variance of the LP coefficient estimates and improving the noise robustness, by manipulating the

power spectrum of the speech signal prior to its approximation by AR model.

III. PROPOSED LP ANALYSIS METHOD

The proposed LP analysis method uses two steps to compute the LP coefficients. In the first step, the manipulation of the periodogram of the input speech is involved in order to reduce the variance of the spectral estimates and remove the parts that are more affected by noise. In the second step, the conventional autocorrelation method is applied using the modified autocorrelation coefficients, which are calculated by taking the inverse FFT of the processed power spectrum. The processed power spectrum is obtained through smoothing followed by thresholding operations. In the smoothing operation, the periodogram of the input speech is smoothed using triangular filters spaced linearly on the Bark frequency scale [11] in order to reduce the spectral estimate variance. This non-linear smoothing, which is inspired by the human auditory system, is performed less at lower frequencies, where the power components are more prominent, and more at higher frequencies, where the power components are less prominent [12], as shown in Fig. 1. Following the smoothing, a thresholding operation is performed, where the influence of low signal-to-noise ratio (SNR) spectral components, which are more affected by noise and add undesirable variance to the spectral estimate, are removed and replaced by the corresponding portions from the smoothed power spectrum, as shown in Fig. 2. As a consequence of the smoothing followed by thresholding operation, the dominant spectral components are preserved, while, the less reliable spectral valleys are ignored and replaced by a smoothed average. Hence, by obtaining the robust power spectrum estimation, its derived-LP coefficients would have better robustness in the presence of noise. The proposed algorithm is described in the following steps:

Step 1: Compute the power spectrum $P(k)$ of a given frame $\{x(n), n = 1, 2, \ldots, N-1\}$ of N samples from a speech signal using the FFT [10]:

$$P(k) = \frac{1}{N} \left| \sum_{n=0}^{M-1} x(n)w(n)e^{-j2\pi kn/M} \right|^2, 0 \leqslant k \leqslant M-1 \tag{8}$$

where $P(k)$ is the estimated power spectrum at the kth normalized frequency bin, M is the size of FFT where $M > N$, and $w(n)$ is a Hamming window.

Step 2: Determine the estimated smoothed power spectrum $\overline{P}(k)$ using a triangular filter at every frequency bin:

$$\overline{P}(k) = \sum_{i=-C(k)}^{C(k)} B(i)P(i-k) \tag{9}$$

where $C(k)$ is half the critical bandwidth of the triangular filter at frequency bin k. The triangular filter $B(i)$ is

which is spaced using the Bark frequency scale, which is given by [11]:

$$Bark(f) = 13 \tan^{-1}(0.00076f) + 3.5 \tan^{-1}\left[\left(\frac{f}{7500}\right)^2\right] \quad (10)$$

Step 3: Using the smoothed $\overline{P}(k)$ as the threshold, $\hat{P}(k)$ is formed by retaining only spectral components that are above the threshold. This is defined as:

$$\hat{P}(k) = \begin{cases} P(k) & if \quad P(k) \geq \overline{P}(k) \\ \overline{P}(k) & otherwise \end{cases} \quad (11)$$

Step 4: Take the inverse FFT of $\hat{P}(k)$ in order to obtain the modified autocorrelation coefficients [10]:

$$\hat{R}(q) = \frac{1}{M}\sum_{k=0}^{M-1}\hat{P}(k)e^{j2\pi kq/M} \quad , \quad 0 \leqslant q \leqslant M-1 \quad (12)$$

These modified autocorrelation coefficients $\hat{R}(q)$, $0 \leq q \leq p$, where p is the LP analysis order, are used in the Levinson-Durbin recursion algorithm [10] to compute the LP coefficients, which we called the Smoothed and Thresholded Linear Prediction (STLP) coefficients.
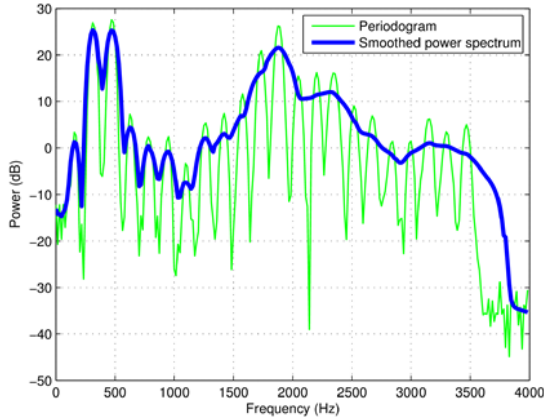

Fig. 1. Periodogram $P(k)$ (thin line) and the smoothed power spectrum $\overline{P}(k)$ (thick line) of speech sound (vowel / e / produced by male speaker).
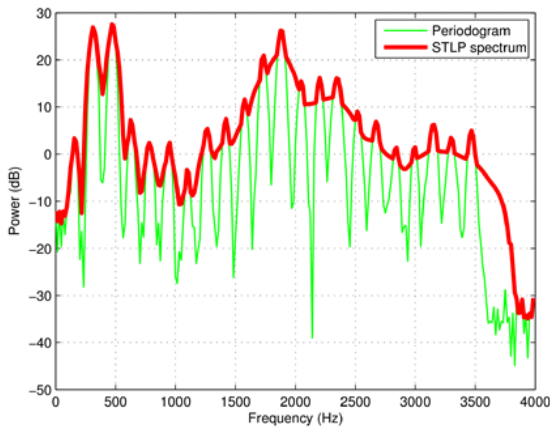

Fig. 2. Periodogram $P(k)$ (thin line) and the resultant spectrum $\hat{P}(k)$ (thick line) after thresholding operation of speech sound (vowel / e / produced by male speaker).

The behaviour of the STLP analysis method in spectral modelling of speech signal is demonstrated in the example shown in Fig. 3. In this figure, the periodogram $P(k)$ of a frame of the speech sound (vowel / e /) is shown together with the all-pole spectral models of order $p = 10$ that were computed with two techniques: the proposed STLP and the AM-LP analysis methods. As displayed in Fig. 3, at high frequencies, the formants attend to have wider bandwidths because of the large critical bandwidths at these frequencies, where more smoothing is performed. Therefore, this added smoothing reduces the influence of noise components at higher frequencies.
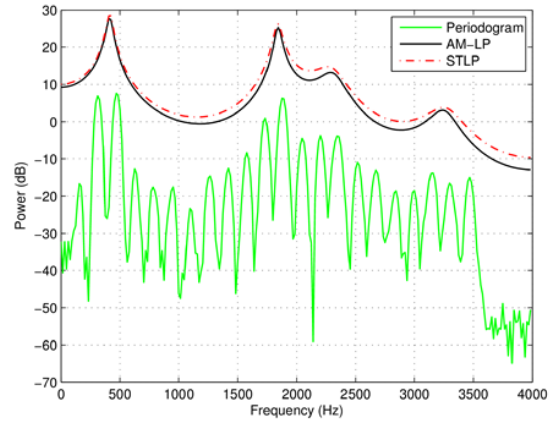

Fig. 3. Periodogram $P(k)$ with corresponding all-pole spectra of order p=10 computed by the conventional LP (solid) and the proposed STLP (dashed) analysis method of clean speech (vowel / e / produced by male speaker). For clarity, the periodogram is moved down by 20 dB.

IV. EXPERIMENTAL SET-UP

A. Database

The TIMIT database [13], which was downsampled to 8 kHz, was utilised for all of the simulations that performed for this paper. The training and testing sets consisted of 3696 sentences produced by 462 speakers of which 136 are female, and 1344 sentences produced by 168 speakers of which 56 are female, respectively. The estimation of the spectral envelope was carried out using the FFT size of 256 frequency samples. A tenth-order of LP analysis with high frequency compensation [1] was performed every 20 ms frame. Thus, the training database consists of 564778 LP parameter vectors, and the evaluation database consists of 206323 LP parameter vectors which were apart of the training database. Because of sharp spectral peaks in the LP spectrum caused by the underestimation of the formant bandwidths, a 10 Hz bandwidth expansion was applied [3].

B. Performance Evaluation Criterion

In order to determine the quality of the power spectrum, the SD of the estimated spectral envelope was computed over the power spectrum of a frequency plane as an objective measure. It is defined as in [3]:

$$SD = \sqrt{\frac{1}{F_s}\int_0^{F_s}\left[10\,log_{10}\,P(f) - 10\,log_{10}\,\hat{P}(f)\right]^2\,df} \quad (13)$$

where F_s is the sampling frequency, and $P(f)$ and $\hat{P}(f)$ are referred to as the true and estimated power spectra, respectively. As can be observed in Equation (13), a low SD indicates the reconstructed speech spectral envelope to be closer to that of the original speech, and therefore is of better quality. This distortion measure is carried out on the power spectrum produced from $20-30$ ms frames of speech. The measure will be utilised to measure the accuracy and robustness of the proposed STLP analysis method in Section V.A.

The number of bits allocated for quantisation influences the performance of the quantisation of the LP parameters. In many cases, the number of bits allocated for quantisation is found when the preferred rate of spectral accuracy has been achieved. This provides an equivalent basis of comparison between the proposed STLP analysis method and the AM-LP, LPSM analysis methods. To evaluate the quantisation process, the SD will be observed in two different classifications: the average SD for the whole data and the percentage of outlier frames. An outlier frame has an SD ≥ 2 dB. The outlier frames are divided into the following types: outlier frames with an SD in the range of $2-4$ dB; and outlier frames with an SD greater than 4 dB. The preferred performance for the quantisation of the LP parameters is when transparent quantisation is achieved [3], which is defined by the following conditions:

- The average SD is about 1 dB.
- The number of outlier frames with an SD between $2-4$ dB is less than 2%.
- No outlier frames are greater than 4 dB.

V. RESULT AND DISCUSSION

A. Noise Robustness Analysis
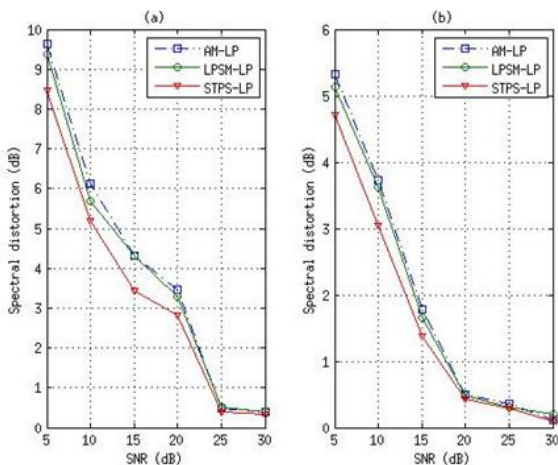

Fig. 4. Spectral distortion values (SD) between the AM-LP, LPSM and STPS spectral envelopes of order p=10 computed from clean and noisy speech (vowel / e / produced by male speaker). Speech was corrupted by two types of noise: (a) additive zero-mean Gaussian white noise, and (b) street noise, in six SNR categories.

The robustness of the spectral envelope estimated using the STLP method was compared with the robustness found using both the AM-LP and LPSM methods. Simulations to find the robustness of the proposed method were carried out by measuring the SD between the power spectrum of the clean and noisy signal, respectively. The results that were obtained from this experiment are shown in Fig. 4. The data shows that the STLP parameters appeared more robust to noise than the AM-LP and LPSM parameters; the SDs of the STLP analysis method were consistently lower than other analysis methods for all SNRs. By referring to the example shown in Fig. 1 and 2, this behavior can be explained by the effectiveness of the smoothing and thresholding operations in the computation of the STLP parameters.

B. Quantisation Performance of LP Parameters

The full-search vector quantisation (VQ) has a high computational complexity that requires excessive memory space in order to perform the quantisation of the codebook. Though the split VQ mechanism is suboptimal, it lowers both the computational complexity and the required memory space to a manageable level without significantly affecting the performance of the VQ [3]. As a result, we used the split VQ to study the quantisation performance of the LP parameters.

In this split VQ, the LSF vector, which is a popular representation of the LP parameters, is separated into different parts of the lower order. The codebooks of the VQ were designed using the Linde-Buzo-Gray (LBG) algorithm [14] with the weighted LSF distance measure for every part, which is given by [3]:

$$d(f,\hat{f}) = \sum_{i=1}^{10}[c_i w_i (f_i - \hat{f}_i)]^2 \qquad (14)$$

where f_i and $\hat{f}_i$ are the ith LSF representation in the approximated and original vector, respectively, and the weights w_i and c_i are assigned to the ith LSF. The variable weight w_i is given by [3]:

$$w_i = [P(f_i)]^r \qquad (15)$$

where $P(f)$ is the LP power spectrum and r is equal to 0.15. The fixed weight c_i is given by [3]:

$$c_i = \begin{cases} 1.0 & for & 1 \leq i \leq 8 \\ 0.8 & for & i = 9 \\ 0.4 & for & i = 10 \end{cases} \qquad (16)$$

We used two-split VQ (where the components of the LSF vector are split into (4,6)) and three-split VQ (where the components of the LSF vector are split into (3,3,4)).

In all experiments, the quantisation performance for each method was evaluated using the SD measure, as given in Equation (13), where $P(f)$ and $\hat{P}(f)$ are the power spectrum of the original and reconstructed speech, respectively, and F_s covers the partial-band SD (i.e. $0-$ 3kHz), which is used to evaluate the quantisation schemes that use a weighted distance measure [15].

The quantisation performances obtained from the two-split VQ experiments are listed in Tables I, II and III for the AM-LP, LPSM and STLP analysis methods, respectively. The results indicate that the two-split VQ that uses the AM-LP and LPSM analysis methods required 24 bits/frame to achieve the transparent quantisation. However, the STLP analysis method required only 22 bits/frame, which saved 2 bits/frame from other techniques. The outlier frames percentage between 2 – 4 dB was much lower in favour of the proposed STLP analysis method.

TABLE I: PARTIAL-BAND AVERAGE SPECTRAL DISTORTION (SD) OF THE TWO-SPLIT VECTOR QUANTISER AS A FUNCTION OF BIT-RATE (USING THE AM-LP ANALYSIS METHOD WITH THE WEIGHTED LSFS DISTANCE MEASURE)

Bits used	Av. SD (in dB)	Outliers (%)	
		2 − 4 dB	> 4 dB
26	0.90	0.39	0.00
25	0.95	0.58	0.00
24	1.04	1.07	0.00
23	1.09	1.68	0.00
22	1.19	3.16	0.00
21	1.26	4.23	0.00
20	1.31	6.10	0.00

TABLE II: PARTIAL-BAND AVERAGE SPECTRAL DISTORTION (SD) OF THE TWO-SPLIT VECTOR QUANTISER AS A FUNCTION OF BIT-RATE (USING THE LPSM ANALYSIS METHOD WITH THE WEIGHTED LSFS DISTANCE MEASURE)

Bits used	Av. SD (in dB)	Outliers (%)	
		2 − 4 dB	> 4 dB
26	0.92	0.87	0.00
25	1.00	1.41	0.00
24	1.03	1.56	0.00
23	1.12	2.67	0.01
22	1.16	2.98	0.03
21	1.26	5.20	0.04
20	1.30	5.95	0.04

TABLE III: PARTIAL-BAND AVERAGE SPECTRAL DISTORTION (SD) OF THE TWO-SPLIT VECTOR QUANTISER AS A FUNCTION OF BIT-RATE (USING THE PROPOSED STLP ANALYSIS METHOD WITH THE WEIGHTED LSFS DISTANCE MEASURE)

Bits used	Av. SD (in dB)	Outliers (%)	
		2 − 4 dB	> 4 dB
26	0.77	0.10	0.00
25	0.84	0.18	0.00
24	0.89	0.23	0.00
23	0.96	0.52	0.00
22	1.01	0.73	0.00
21	1.11	1.55	0.00
20	1.18	2.30	0.00

TABLE IV: PARTIAL-BAND AVERAGE SD OF THE THREE-SPLIT VECTOR QUANTISER AS A FUNCTION OF BIT-RATE (USING THE AM-LP ANALYSIS METHOD WITH THE WEIGHTED LSFS DISTANCE MEASURE)

Bits used	Av. SD (in dB)	Outliers (%)	
		2 − 4 dB	> 4 dB
30	0.78	0.20	0.00
29	0.80	0.27	0.00
28	0.85	0.51	0.00
27	0.88	0.56	0.00
26	0.97	0.91	0.00
25	1.05	1.73	0.00
24	1.18	3.10	0.01
23	1.21	3.83	0.01

Tables IV, V and VI illustrate the results obtained from the three-split VQ for the AM-LP, LPSM and STLP analysis methods, respectively. It can be seen that the proposed STLP analysis method offers an advantage of 1 and 2 bits/frame over the AM-LP and LPSM analysis methods, which required 25 and 26 bits/frame to achieve the quantisation transparency, respectively.

TABLE V: PARTIAL-BAND AVERAGE SD OF THE THREE-SPLIT VECTOR QUANTISER AS A FUNCTION OF BIT-RATE (USING THE LPSM ANALYSIS METHOD WITH THE WEIGHTED LSFS DISTANCE MEASURE)

Bits used	Av. SD (in dB)	Outliers (%)	
		2 − 4 dB	> 4 dB
30	0.79	0.42	0.00
29	0.86	0.72	0.00
28	0.92	1.00	0.00
27	0.95	1.09	0.00
26	1.01	1.61	0.00
25	1.04	2.52	0.01
24	1.19	3.65	0.03
23	1.21	4.83	0.04

TABLE VI: PARTIAL-BAND AVERAGE SD OF THE THREE-SPLIT VECTOR QUANTISER AS A FUNCTION OF BIT-RATE (USING THE PROPOSED STLP ANALYSIS METHOD WITH THE WEIGHTED LSFS DISTANCE MEASURE)

Bits used	Av. SD (in dB)	Outliers (%)	
		2 − 4 dB	> 4 dB
30	0.66	0.09	0.00
29	0.68	0.11	0.00
28	0.77	0.33	0.00
27	0.82	0.40	0.00
26	0.85	0.45	0.00
25	0.96	1.20	0.00
24	1.01	1.36	0.00
23	1.09	1.76	0.00

VI. CONCLUSION

This paper presented a modified method for estimating LP coefficients for current speech coders, by applying nonlinear smoothing followed by thresholding to the power spectrum. These STLP coefficients possess improved robustness to noise. The advantage of this method is that it is fully compatible with current speech coder implementations. The LP quantisation performance of the STLP coefficients, in comparison with the AM-LP and LPSM coefficients, was investigated for various split vector quantisation schemes. For two-split vector quantisation, the proposed STLP analysis method offered a saving of 2 bits/frame over the aforementioned methods. Additionally, for three-split vector quantisation, the proposed method offered savings of 1 and 2 bits/frame over the AM-LP and LPSM methods, respectively, in terms of the spectral distortion measure. The results demonstrated the improved noise-robustness of the STLP coefficients for low to medium SNR levels and for both white and street noise. Since the Bark frequency triangular filters used for the power spectrum smoothing are similar to those used in the computation of Mel frequency cepstral coefficients (or MFCCs), these STLP coefficients may exhibit better recognition performance and noise robustness in automatic speech recognition

tasks. We will investigate the use of STLP coefficients in the network speech recognition context in a future paper.

REFERENCES

[1] J. Makhoul, "Linear prediction: A tutorial review," in *Proc. the IEEE,* 1975, pp. 561-580.

[2] B. S. Atal, V. Cuperman, and A. Gersho, *Speech and Audio Coding for Wireless and Network Applications*, Springer Science & Business Media, 2012, pp. 191-201.

[3] W. B. Kleijn and K. K. Paliwal, *Speech coding and synthesis.* Elsevier Science Inc., New York, NY, USA, 1995, pp. 443-466.

[4] S. M. Kay, "The effects of noise on the autoregressive spectral estimator," *IEEE Transactions on Acoustics, Speech and Signal Processing,* vol. 27, no. 5, pp. 478-485, 1979.

[5] M. R. Sambur and N. S. Jayant, "Lpc analysis/synthesis from speech inputs containing quantizing noise or additive white noise," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 24, no. 6, pp. 488-494, 1976.

[6] A. Trabelsi, F. Boyer, Y. Savaria, and M. Boukadoum, "Improving lpc analysis of speech in additive noise," in *Proc. IEEE Northeast Workshop on Circuits and Systems*, 2007, pp. 93-96.

[7] Y. Kuroiwa and T. Shimamura, "An improvement of lpc based on noise reduction using pitch synchronous addition," in *Proc. IEEE International Symposium on Circuits and Systems*, 1999, pp. 122-125.

[8] Y. Liang, C. Bao, B. Xia, N. Li, and R. Li, "An lp spectrum modification method for noisy speech based on linear extrapolation," in *Proc. 35th International Conference on Telecommunications and Signal Processing (TSP)*, 2012, pp. 450-453.

[9] D. B. Paul, "The spectral envelope estimation vocoder," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 29, no. 4, pp. 786-794, 1981.

[10] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*, John Wiley & Sons, 2009, pp. 124-202.

[11] H. Fletcher, "Auditory patterns," *Reviews of Modern Physics*, vol. 12, no. 1, pp. 47-65, 1940.

[12] B. Moore, *An Introduction to the Psychology of Hearing*, 6th ed. Academic Press, 2012, ch. 1.

[13] J. S. Garofolo, L. F. Lamel, W. M. Fisher, J. G. Fiscus, D. S. Pallett, and N. L. Dahlgren, *TIMIT Acoustic Phonetic Continuous Speech Corpus*, Philadelphia; Linguistic Data Consortium, 1993.

[14] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Transactions on Communications,* vol. 28, no. 1, pp. 84-95, 1980.

[15] S. So and K. K. Paliwal, "Efficient product code vector quantisation using the switched split vector quantiser," *Digital Signal Processing*, vol. 17, no. 1, pp. 138-171, 2007.

Aadel Alatwi was born in Tabouk, Saudi Arabia, in 1980. He received the B.S. degree from King Abdul-Aziz University, Jeddah, Saudi Arabia, in 2004, the M.S. degree from Griffith University, Brisbane, Australia, in 2008, both in communication engineering. He is currently a PhD candidate in the School of Engineering at Griffith University, Brisbane, Australia. His research interests include speech coding and speech recognition.

Kuldip K. Paliwal was born in Aligarh, India, in 1952. He received the B.S. degree from Agra University, Agra, India, in 1969, the M.S. degree from Aligarh Muslim University, Aligarh, India, in 1971, and the PhD degree from Bombay University, Bombay, India, in 1978. He is a professor in the School of Engineering at Griffith University, Brisbane, Australia. His current research interests include speech coding, speech and speaker recognition, speech enhancement, face recognition, image coding, pattern recognition and artificial neural networks.