WIDE-BAND SPEECH CODING USING KERNEL METHODS AND BANDWIDTH EXTENSION BASED ON PARAMETRIC STEREO

Gh. Alipoor[1] and M.H. Savoji[2]

Electrical and Computer Engineering Faculty
Shahid Beheshti University, Evin Square, Tehran 1983963113, Iran
[1]g-alipoor@sbu.ac.ir [2]m-savoji@sbu.ac.ir

ABSTRACT

A novel wide-band speech coding scheme is developed, in this paper, based on kernel methods and bandwidth extension. The KLMS algorithm, a kernelized version of the well-known LMS algorithm, is employed in the framework of the backward ADPCM technique for encoding the narrow-band part of the wide-band speech. Simulation results show that utilizing this nonlinear method results in an average improvement of up to 3.4 dB in the SNR and 0.28 in the PESQ measure of the decoded speech. The resultant narrow-band codec is subsequently extended to the wide-band speech using a novel bandwidth extension technique inspired by the parametric stereo coding. It is shown that the KLMS algorithm is also effective in this framework. This leads to a wide-band speech coding scheme built on the nonlinear narrow-band codec at the extra cost of a small increase in the bite rate.

Index Terms— Bandwidth extension, kernel least mean square, kernel methods, parametric stereo, wide-band speech coding

1. INTRODUCTION

There is an ever increasing demand for wide-band (50-7000 Hz) speech with the growing popularity of various applications such as teleconferencing, IP-telephony, wireless communication and multimedia services. Despite the obvious advantages of wide-band coding techniques and their developments in terms of several algorithms and standards, narrow-band (200-3400 Hz) speech is still more widely used. This is mainly due to the narrow-band nature of the existing mobile communication systems and landline public switched telephone service networks (PSTN). Therefore, a key requisite of wide-band coding techniques is their interoperability with existing narrow-band standards and already installed infrastructures. However, new wide-band coding techniques are developed with a view of their compatibility with future super-wide-band [1] and full-band [2] coding applications.

Bandwidth extension (BWE) [3] is a technique to extend narrow-band speech coding algorithms to wide-band speech at the cost of a small increase in the bit rate. The main idea is based on the fact that there is a noticeable correlation between the low and high frequency parts of the speech signal. This means that it is possible to reproduce, at the decoder, the high-band (HB) signal from the low-band part without the need for transmitting the whole information pertaining to the HB signal. This technique operates on the top of the encoded narrow-band speech and extents the coding algorithm to the whole wide band. Apart from interoperability, this hierarchical scheme ensures bandwidth and bit-rate scalability, features that are desirable in emerging applications.

BWE techniques are mostly based on the source-filter model, in which the HB signal is restored at the decoder by readjusting the spectrum of a locally generated HB excitation signal with a spectral shaping filter. This spectral shape can be estimated, at the decoder, based on linear prediction (LP) analysis of the decoded low-band speech. But, normally, a relatively higher performance is expected if the LP filter is extracted, at the encoder, based on the original HB signal and transmitted as side information. This technique is successfully utilized in the ITU g729.1 standard, as well as the ITU g722.2 (AMR-WB) and AMR-WB+ codecs. Another relevant method is the powerful but complicated spectral band replication (SBR) technique which is also used and standardized in MPEG-4 high-efficiency AAC (HE-AAC) [4]. On the other hand, inspired by the structure proposed in [5] for parametric stereo (PS) coding, a novel SBE technique is devised in [6] in which the spectral contents of the HB signal are directly reconstructed based on the received side information. In addition to its good performance, this extension technique avoids the necessity for LP analysis and reveals higher bit-rate scalability.

The core part of the BWE technique is an efficient coding scheme for the narrow-band speech. Adaptive differential pulse code modulation (ADPCM) is a technique

This work is partially supported by the Iranian Research Institute for IC-ITRC.

vastly used for this purpose [7]. In this technique the quantization is carried out on the residual signal or what remains of the speech signal when its predictable parts have been removed adaptively. Linear prediction is the simplest choice in this paradigm where prediction is performed by a linear combination of a finite number of past samples. Despite their simplicity in conception and implementation, linear models are unable to account for nonlinear characteristics inherent in speech signals. On the other hand, our previously reported study shows the usefulness of nonlinear processing based on emerging methods of kernel adaptive filtering in this context [8].

We examine here employing the kernel LMS (KLMS) algorithm, which is judged as the best kernel adaptive filtering algorithm based on our previous study [8], in the framework of the ADPCM technique for narrow-band speech coding. The resultant codec is then extended to the wide-band speech, to address the need for wide-band speech coding, using the PS-based bandwidth extension technique. The paper is organized as follows. The KLMS algorithm is briefly described in section 2 following a general introduction on kernel methods. Employing this algorithm within the ADPCM technique for coding narrow-band speech and its extension to wide band are addressed in sections 3 and 4, respectively. Section 5 is dedicated to simulation results. Finally some conclusion remarks are presented in section 6.

2. KERNEL METHODS AND KLMS ALGORITHM

A Hilbert space $\mathcal{H}$ over a set $\mathcal{X}$ is called a reproducing kernel Hilbert space [9] if there exists a (real or complex valued) function $K: \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ for which:

i) Every function defined as $k_x(\cdot) =: K(\boldsymbol{x},\cdot)$ belongs to $\mathcal{H}$.
ii) The so defined function k_x satisfies the reproducing property, i.e. $for\ \forall \boldsymbol{x} \in \mathcal{X}, \forall f \in \mathcal{H};\quad f(\boldsymbol{x}) = \langle f, k_x \rangle_{\mathcal{H}}$.

It can be shown that for any RKHS $\mathcal{H}$, one can imagine a space, known as the feature space, in which the inner product can be calculated through evaluating its kernel function K in the input space. This mapping, denoted by ϕ and termed feature mapping, projects the input $\boldsymbol{x} \in \mathcal{X}$ as the function $\phi(\boldsymbol{x})(\cdot) = K(\boldsymbol{x},\cdot) \in \mathcal{H}$. In other words, representing the function $\phi(\boldsymbol{x})(\cdot)$ as $\phi(\boldsymbol{x})$, the kernel K corresponds to a feature mapping ϕ for which

$$K(\boldsymbol{x}, \boldsymbol{y}) = \langle \phi(\boldsymbol{x}), \phi(\boldsymbol{y}) \rangle \tag{1}$$

Equation (1) is known as the kernel trick and states that the inner product in the feature space can be expressed in terms of the kernel function evaluation. Kernel trick has the central role in kernel methods based on which all linear inner-product-based algorithms can be implicitly applied to the feature space while remaining in the input space. The real value of the kernel trick is realized when one considers that in practice feature spaces are usually of high, or even infinite, dimensions. Therefore, one can implicitly extend linear algorithms, such as those used in optimization problems, to a high-dimensional feature space while performing all calculations in the low-dimensional input space. The linear optimization algorithms developed in RKHS possess the properties of convexity and universal nonlinear approximation. Furthermore, nonlinear kernel methods are quite flexible so that one can change the nonlinear model just by changing the kernel function used.

Extending linear adaptive algorithms to RKHSs are mostly based on reformulating original algorithms in terms of inner products and then replacing the inner products with the kernel function evaluations. This will be equivalent to implicitly solving the linear adaptive algorithms in feature spaces induced by kernel functions, where transformed signals are more likely to be linearly related to the so called desired signal. Kernel methods are mostly derived in batch mode that involve Gram matrix whose dimensionality is commensurate with the number of data points. Hence these methods, as such, are excluded in most online applications. The solution to this problem has been addressed in numerous online kernel methods [10]. The milestone in the evolution of kernel adaptive algorithms is the KLMS algorithm which is a straightforward extension of the linear least mean square (LMS) algorithm into RKHS [11].

In the framework of ADPCM speech coding with backward prediction, we aim at predicting the current speech sample $s(i)$ based on P past samples of the reconstructed speech $\hat{s}$. Using the normalized LMS (NLMS) algorithm, the weight update equation, at instant i, is:

$$\boldsymbol{w}_i = \boldsymbol{w}_{i-1} + \frac{\mu \boldsymbol{x}_i \hat{e}(i)}{\tilde{\sigma}_i^2} \tag{2}$$

$\boldsymbol{x}_i = [\hat{s}(i-1) \quad \cdots \quad \hat{s}(i-P)]^T$ and $\hat{e}(i)$ are the input vector and the quantized value of the prediction error at instant i, respectively. $0 < \mu \ll 1$ is the convergence parameter to control the memory span of the predictor filter and therefore the convergence speed of the algorithm and $\tilde{\sigma}_i^2$ is an estimate of the input signal variance.

The KLMS algorithm [10-11] is derived by employing the NLMS algorithm to predict $s(i)$ based on the transformed input $\boldsymbol{\varphi}_i = \phi(\boldsymbol{x}_i)$. Denoting by $\boldsymbol{\omega}$ the estimated value of the filtering coefficients in the feature space and assuming $\boldsymbol{\omega}_0 = \boldsymbol{0}$, it is easily seen that:

$$\tilde{s}(i) = \mu \sum_{j=1}^{i-1} \frac{\hat{e}(j)}{\tilde{\sigma}_j^2} K(\boldsymbol{x}_j, \boldsymbol{x}_i) \tag{3}$$

$$\tilde{\sigma}_j^2 = \alpha \tilde{\sigma}_{j-1}^2 + (1-\alpha) K(\boldsymbol{x}_j, \boldsymbol{x}_j) \tag{4}$$

$\tilde{\sigma}_j^2$ is an estimate of the variance of the transformed data at instant j and α is the forgetting factor in this estimation.

Therefore, at instant i, the coefficient estimate can be expressed as a linear combination of all the previous and present transformed inputs, weighted by the prediction errors and scaled by the convergence parameter μ. At instant i, the predicted signal, i.e. the filter output, is:

$$\tilde{s}(i) = \boldsymbol{\omega}_{i-1}^T \boldsymbol{\varphi}_i = \left[\mu \sum_{j=1}^{i-1} \hat{e}(j) \boldsymbol{\varphi}_j^T\right] \frac{\boldsymbol{\varphi}_i}{\tilde{\sigma}_j^2} \tag{5}$$

which is efficiently computed using the kernel trick in the input space, as:

$$\tilde{s}(i) = \mu \sum_{j=1}^{i-1} \frac{\hat{e}(j)}{\tilde{\sigma}_j^2} K(\boldsymbol{x}_j, \boldsymbol{x}_i) \quad (6)$$

In conclusion, adaptive NLMS filtering can be implicitly carried out in the high-dimensional feature space without direct access to the feature map and the filtering coefficients. More interestingly, it has been shown that the KLMS algorithm possesses the property of self-regularization that makes an extra regularization unnecessary [11]. In addition to simplifying the implementation, this property improves the performance because regularization biases the optimal solution.

As one can see, the size of the network over which the signal is expanded or the number of past samples based on which the signal is estimated, called dictionary, increases with the size of the data. Alleviating this problem is the main implementation challenge in online applications where the number of observations continuously increases. In practice, redundancy among input data makes it possible to drastically reduce the size of the estimation memory, at the cost of a negligible effect on the quality of the model. This is generally carried out based on selecting the most informative data and discarding the others from the dictionary. This procedure is termed sparsification and many techniques have been proposed for this purpose in both batch and online modes. One of the first and still widely used measures is the novelty criterion (NC) proposed in [12] which acts based on a simple distance measure in the input space. In this technique, at iteration i, the minimum distance of the new input vector $\boldsymbol{x}_i$ to all the vectors retained in the dictionary $\mathcal{C}_{i-1}$ (i.e. $\min_{\boldsymbol{x}_j \in \mathcal{C}_{i-1}} \|\boldsymbol{x}_i - \boldsymbol{x}_j\|$) is calculated. The new input vector will be accepted as a new element of the dictionary only if this measure is larger than a preset threshold, and the quantized prediction residual $\hat{e}(i)$ is also larger than another preset threshold. This sparsification drastically reduces the complexity of the online algorithm. This in turn makes kernel adaptive filtering a competitive candidate for nonlinear adaptive signal processing.

3. NARROW-BAND SPEECH CODING BASED ON KERNEL METHODS

The main source of performance improvement for the ADPCM coder is the reduced dynamic range of the quantizer input signal. This reduction is achieved by removing the short term redundancy of the speech waveform accomplished by subtracting an adaptively predicted signal from the input signal. In backward prediction, used in this work, coding parameters are sequentially estimated from the past quantized residual signal, also available at the decoder. Prediction is usually performed linearly. But, speech is inherently nonlinear and nonlinear filters with higher ability to cope with nonlinearity ought to be used. Using nonlinear prediction a higher prediction gain is expected. Volterra filters are nonlinear models widely used for this purpose [13-14]. At instant i utilizing a Volterra filter consisting of a linear term of memory size P_1 with coefficients $\boldsymbol{\omega}_{1_i}(k)$ and a quadratic term of memory size P_2 with coefficients $\boldsymbol{\omega}_{2_i}(k_1, k_2)$, the backward ADPCM predicted value is calculated as:

$$\begin{aligned}\tilde{s}(i) = {} & \sum_{k=1}^{P_1} \boldsymbol{\omega}_{1_i}(k)\hat{s}(i-k) + \\ & \sum_{k_1=1}^{P_2} \sum_{k_2=k_1}^{P_2} \boldsymbol{\omega}_{2_i}(k_1, k_2)\hat{s}(i-k_1)\hat{s}(i-k_2)\end{aligned} \quad (7)$$

However, in addition to their inherent instability, the fact that their computational complexity grows exponentially with the memory size and the degree of the nonlinearity involved is the major obstacle for their practical use.

Nonlinear adaptive Volterra filtering can be also accomplished using kernel adaptive algorithms. To find a proper kernel function corresponding to the Volterra filter depicted in (7) and assuming P_1=P_2=P, we define a mapping from $\mathbb{R}^P$ to $\mathbb{F} = \mathbb{R}^{P+P(P+1)/2}$ as:

$$\begin{aligned}\phi(\boldsymbol{x}_i) = {} & [\hat{s}(i-1), \cdots, \hat{s}(i-P), \hat{s}(i-1)\hat{s}(i-1), \\ & \sqrt{2}\,\hat{s}(i-1)\hat{s}(i-2), \cdots, \sqrt{2}\,\hat{s}(i-1)\hat{s}(i-P), \\ & \sqrt{2}\,\hat{s}(i-2)\hat{s}(i-1), \hat{s}(i-2)\hat{s}(i-2), \cdots, \\ & \sqrt{2}\,\hat{s}(i-2)\hat{s}(i-P), \cdots, \hat{s}(i-P)\hat{s}(i-P)]^T\end{aligned} \quad (8)$$

The mapping $\phi(\cdot)$ transforms the input vector $\boldsymbol{x}_i \in \mathbb{R}^P$ to a vector containing all possible first and second order permutations of the elements of $\boldsymbol{x}_i$. Defining $\boldsymbol{\omega}_i$ as $[\boldsymbol{\omega}_{1_i}; \boldsymbol{\omega}_{2_i}]$, it is easy to see that $\tilde{s}(i) = \boldsymbol{\omega}_i{}^T \phi(\boldsymbol{x}_i)$. Furthermore, one can show that:

$$\phi(\boldsymbol{x}_i)^T \phi(\boldsymbol{x}_j) = (\boldsymbol{x}_i{}^T \boldsymbol{x}_j) + (\boldsymbol{x}_i{}^T \boldsymbol{x}_j)^2 =: K(\boldsymbol{x}_i, \boldsymbol{x}_j) \quad (9)$$

Fortunately, mapping function ϕ constitutes an RKHS whose kernel function is defined above [15]. In contrast to the Volterra filter, the estimation complexity is now linearly dependent on the input dimensionality.

4. WIDE-BAND SPEECH CODING BASED ON PARAMETRIC STEREO

A similar idea to BWE is used in the parametric stereo audio coding technique proposed in [5] where the substantial redundancy existent between the two channels is exploited. The stereo audio signal is encoded, in this technique, as a monaural signal, made by linear combination of two channels in the frequency domain, plus a small amount of parametric overhead to extract the stereo channels. The needed parameters are calculated and transmitted. They are retrieved in the decoder. Their calculation, encoding and retrieval are based on spatial psychoacoustical principles. The monaural signal, also known as the downmix signal, can be encoded using any (conventional) audio coder.

A novel BWE method is devised in [6], based on the above mentioned PS technique, to efficiently encode the monaural wide-band speech signal. In this way, the wide-band signal is split, at the encoder, into low and high bands and the LB signal is treated similar to downmix signal and therefore encoded using a core codec developed for narrow-band speech. The parameters describing the HB part are extracted based on the spectral contents of the original HB and the decoded LB signals. The two signals are first segmented by an analysis windowing process and subsequently transformed to frequency domain. Each transformed segment is divided into non-overlapping sub-bands and used to extract the necessary parameters. The bandwidths of the sub-bands are chosen in accordance with the frequency resolution of the human auditory system. Three parameters are extracted, for each frequency band, based on the frequency contents of the corresponding sub-bands of LB and HB signals. These parameters are the inter-channel intensity difference (IID) that characterizes the power of the LB signal in proportion to that of the HB part, the inter-channel phase difference (IPD), to represent the phase difference between the LB and HB signals, and the inter-channel coherence (IC) defined as the normalized cross-correlation coefficient. All these parameters are subsequently quantized based on a perceptual criterion and sent to the decoder along with the encoded low-band signal. The high-frequency part is regenerated, at the decoder, using the decoded LB part with the aid of the received parameters. Finally, wide-band speech is reconstructed from the retained low and high-band signals by sub-band composition. This method makes it possible to extend the narrow-band codec to wide-band speech with a small increase in the bit rate. Further details are found in [6].

5. RESULTS

All results reported in this paper are averaged over 48 speech signals, 24 male and 24 female speakers, selected from the test set of the DARPA TIMIT database [16]. These signals, each containing a whole sentence in English, are originally sampled at 16 kHz. Narrow-band speech signals, when used in isolation, are generated by down-sampling the wide-band signals to 8 kHz, after applying a 20th order anti-aliasing low-pass filter. Both narrow and wide-band signals are quantized uniformly at 16 bits/sample. Moreover, the quality evaluation of reconstructed speech signals is conducted based on signal to noise ratio (SNR) and perceptual evaluation of speech quality (PESQ) criteria. PESQ evaluation of narrow and wide-band speech is done as suggested by ITU-T P.862 [17] and ITU-T P.862.2 [18] recommendations, respectively. These standards recommend procedures based on which to evaluate the subjective quality of the reconstructed (narrow and wide-band) speech in comparison to their original counterparts.

The averaged results, utilizing normalized LMS and KLMS adaptive algorithms with memory span $P = 10$ in backward ADPCM speech coding, are summarized in Table 1. Four bit-per-sample (bps) values are used for scalar quantization of the residual. Results of the KLMS algorithm are achieved using novelty criterion sparsification technique to control the memory size. It can be seen that utilizing the KLMS adaptive algorithm, as compared to the NLMS algorithm, results in average improvements of up to 0.28 in the PESQ measure and up to 3.4 dB in the reconstruction SNR. This improvement is also confirmed by our informal listening tests. This is while results reported in [14] shows that using quadratic Volterra filter, by itself, does not cause a significant improvement in the coding performance. The improvement achieved by the KLMS algorithm is even higher than the improvement achieved using the computationally expensive schemes devised in [14] for optimum utilization of nonlinear Volterra filtering. Furthermore, remaining in the lower-dimensional input space and utilizing the sparsification techniques makes it possible to implement nonlinear kernel adaptive filtering algorithms with a moderate complexity. The complexity is very low as compared to the Volterra filtering algorithms equipped with the schemes developed for optimum utilization of nonlinear prediction. Specifically, two optimum Volterra-based schemes examined in [14] take, on average, 46.6 and 237 seconds processing time, respectively. But, the average processing time of the KLMS algorithm is 10.9 seconds. This is in comparison with 5.3 seconds for the LMS algorithm. All these tests are done on a same machine. In fact, the selected polynomial kernel implements the quadratic Volterra filter exactly, while implementing this adaptive filter in the lower-dimensional input space avoids some instability characteristics the Volterra filters suffer from. Moreover, the self-regularization property of the KLMS algorithm results in more reliable numerical solutions in the sense that the algorithm is less sensitive to numerical accuracy.

As for the wide-band speech codec, the average performance is summarized in Table 2, using the PS-based BWE method. In these tests sub-band decomposition of the wide-band speech is carried out using wavelet transform and the LB signal is encoded using the backward ADPCM with either normalized LMS or KLMS algorithms. Parameters describing the HB part are updated at the rate of 20 frames per second where each frame is divided into 5 sub-bands. Moreover, IID, IPD and IC sub-band parameters are quantized, based on the method proposed in [5], using 5, 3 and 2 bits, respectively. As can be seen, kernelizing the LMS adaptive filtering algorithm improves the overall performance of the wide-band codec. This improvement is up to 0.08 in the PESQ measure and up to 0.7 dB in the reconstruction SNR. Furthermore, utilizing bandwidth extension technique makes it possible to extent the narrow-

Table 1 Results of narrow-band speech coding

	bps=3	bps=4	bps=5	bps=6
(a) **LMS**				
PESQ	3.5625	3.9530	3.9530	4.3288
SNR (dB)	17.3006	21.3992	25.2355	28.9038
(b) **KLMS**				
PESQ	3.5586	3.9773	4.2338	4.3738
SNR (dB)	19.2107	24.0261	28.2057	32.3236

Table 2 Results of wide-band speech coding

	bps=3	bps=4	bps=5	bps=6
(a) **LMS**				
PESQ	3.5134	3.9088	4.1380	4.2822
SNR (dB)	12.1517	13.3230	14.0278	14.4345
(b) **KLMS**				
PESQ	3.5398	3.9735	4.2216	4.3546
SNR (dB)	12.8200	13.9954	14.4931	14.7338

band core codec to wide-band speech. This extension is achieved at the cost of a relatively small increase of 1 kbps in the bit rate. One can compare this scheme, from bit-rate and quality points of view with the ITU g722.2 (AMR-WB) standard in which bandwidth extension is utilized in encoding the high-band signal and the ITU g729.1 standard where the high band is coded separately in the DCT domain.

6. CONCLUSION

In response to the ever increasing demand for high-quality speech, a novel wide-band speech coding scheme is proposed in this paper and its performance examined. The utilization of kernel methods, in the framework of the backward ADPCM technique, for narrow-band speech coding is studied for the first time. Simulation results show that utilizing the well-known KLMS algorithm in ADPCM speech coding results in a considerable improvement in the overall quality of the decoded speech. This improvement is up to 3.4 dB and 0.28 increase in the average SNR and PESQ measures, respectively. These results are considerable as compared to that of the nonlinear Volterra filters. In fact, the selected polynomial kernel implements the quadratic Volterra filter exactly; while implementing it in the lower-dimensional input space avoids some instability characteristics the Volterra filters suffer from. Furthermore, the self-regularization property of the KLMS algorithm results in more reliable numerical solutions.

This scheme is subsequently extended to the wide-band speech using a bandwidth extension technique based on parametric audio stereo coding. The adopted BWE method makes use of the noticeable correlation that exists between the low and high frequency parts of the speech signal. It is shown that it is possible to satisfactorily regenerate the higher-band signal based on the lower frequency components with the aid of a small amount of side information. This in turn leads to a wide-band speech coding algorithm at the extra cost of a small increase in the overall bit rate. This technique is however potentially interesting also for extending the wide-band speech to super-wide band or even full band. Our tests showed the good robustness of the codec against background noise, however, the effect of transmission noise on this backward adaptive scheme needs more investigations. Furthermore, it would be of importance to conduct standard subjective assessments of the resultant wide-band coodec. All these topics constitute our future line of work.

7. REFERENCES

[1] Lasse Laaksonen, *et al.*, "Superwideband Extension of G.718 and G.729.1 Speech Codecs," in *Interspeech*, Tokyo, Japan, 2010.

[2] A. Taleb and S. Karapetkov, "G.719: The first ITU-T standard for high-quality conversational fullband audio coding," *Communications Magazine, IEEE,* vol. 47, pp. 124-130, 2009.

[3] P. Jax and P. Vary, "Bandwidth extension of speech signals: a catalyst for the introduction of wideband speech coding?," *Communications Magazine, IEEE,* vol. 44, pp. 106-111, 2006.

[4] B. Geiser, *et al.*, "Bandwidth Extension for Hierarchical Speech and Audio Coding in ITU-T Rec. G.729.1," *Audio, Speech, and Language Processing, IEEE Transactions on,* vol. 15, pp. 2496-2509, 2007.

[5] J. Breebaart, *et al.*, "Parametric Coding of Stereo Audio," *EURASIP Journal on Applied Signal Processing,* vol. 9, pp. 1305-1322, 2005.

[6] M. Ghaderi and M. H. Savoji, "Wideband speech and audio coding using a new spectral replication method based on parametric stereo coding," in *Intelligent Signal Processing and Communication Systems (ISPACS), 2010 International Symposium on*, 2010, pp. 1-4.

[7] N. S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*. New Jersey: Prentice Hall, 1984.

[8] G. Alipoor and M. H. Savoji, "Nonlinear Backward ADPCM Speech Coding Using Kernel Methods and Online Sparsification," *Accepted for Publication in Transactions on Emerging Telecommunications Technologies,* 2012.

[9] B. Schölkopf and A. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond* Cambridge, MA: MIT Press, 2002.

[10] W. Liu, *et al.*, *Kernel Adaptive Filtering: A Comprehensive Introduction*. Hoboken, New Jersey: John Wiley & Sons, Inc., 2010.

[11] W. Liu, *et al.*, "The Kernel Least Mean Square Algorithm," *IEEE Transactions on Signal Processing,* vol. 56, pp. 543–554, 2008.

[12] J. Platt, "A Resource Allocating Network for Function Interpolation," *Neural Computation,* vol. 3, pp. 213–225, 1991.

[13] G. L. Sicuranza, "Quadratic Filters for Signal Processing.," *IEEE Proceedings,* vol. 80, pp. 1263-1285, 1992.

[14] G. Alipoor and M. H. Savoji, "Employing Volterra Filters in the ADPCM Technique for Speech Coding: A Comprehensive Investigation," *European Transactions on Telecommunications,* vol. 22, pp. 81-92, 2011.

[15] M. O. Franz and B. Schölkopf, "A Unifying View of Wiener and Volterra Theory and Polynomial Kernel Regression," *Neural Computation,* vol. 18, pp. 3097-3118, 2006.

[16] TIMIT. "DARPA TIMIT-Acoustic-Phonetic Continuous Speech Corpus," [Online].

[17] ITU-T, "P.862: Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs," in *ITU-T Recommendation*, ed. Geneva, Switzerland, 2001.

[18] ITU-T, "P.862.2: Wideband extension to recommendation for the assessment of wideband telephone networks and speech codecs," in *ITU-T Recommendation*, ed. Geneva, Switzerland, 2005.