Performance of Immittance Spectral Pairs (ISP) in LPC Scalar Quantizer Schemes

Shlomo Peller and Yuval Bistritz

Department of Electrical Engineering - Systems, Tel Aviv University, Tel Aviv 69978, Israel

Abstract. Immittance Spectral Pairs (ISP) was recently proposed as an alternative to LSP for parametrizing an LPC filter. The ISP parametrization for an LPC filter of order p consist of $p-1$ 'frequency' parameters and a 'gain' parameter, as opposed to p 'frequencies' of the LSP. ISP shares with LSP the ordering property for the 'frequencies' and, similarly, it comes with necessary and sufficient conditions for the stability of the LPC filter. Preliminary experimental results of LPC quantization showed that ISP is a potential favorable competitor to LSP. This paper presents the results of an experimental comparative study of the performance of ISP and LSP in several coding schemes. The results consistently show advantage of ISP over LSP with no increase in computational complexity.

1. Introduction

Since Line Spectra Pairs (LSP) has been proposed for speech encoding [1], it has been studied by many researchers, and has become the most common set of parameters for quantizing and encoding the spectral envelope of speech signals.

Recently, we introduced an alternative to LSP, called Immittance Spectral Pairs (ISP) [2]. The ISP parametrization for an LPC filter of order p consist of $p-1$ 'frequency' parameters and a 'gain' parameter, as opposed to the p 'frequencies' of LSP. ISP shares with LSP the ordering property for the frequencies and, similarly, it comes with necessary and sufficient conditions for the LPC filter's stability. In the preliminary experimental results of LPC quantization reported in [2], ISP was found to compare favorably with LSP. On a theoretical basis that views LPC as a model of the vocal tract, it may be argued that ISP, unlike LSP, does not pose artificial boundary conditions at the glottis or the lips [3]. Consequently, a resulting closer tracking of formants might be expected to render better quantization performance.

This paper brings the results of a comparative experimental study of the performance of ISP and LSP in several coding schemes. In our experiments we implemented, for both LSP and ISP, some of the more successful quantization schemes which have been proposed for LSP during the last decade. The results consistently show a relative advantage in favor of ISP. These findings may suggest ISP as a worthy substitute for LSP in LPC coders, with the prospect of improving bit-rate to distortion characteristics with no increase (but actually some decrease) in computational complexity.

2. ISP parameters

In the realization of the LPC filter $1/A_p(z)$ by a cascade of p lattice sections that implements the Levinson recursion for $n = 1, \ldots, p$

$$\begin{aligned} A_n(z) &= A_{n-1}(z) - k_n z^{-1} A^{\#}_{n-1}(z) \qquad (1) \\ A^{\#}_n(z) &:= z^{-n} A_n(z^{-1}), \end{aligned}$$

the LSP representation corresponds to two extreme boundary conditions, obtained by extending the recursions - once with $k_{p+1} = +1$, and once with $k_{p+1} = -1$. In the pseudo model for the vocal tract, the two LSP polynomials represent two 'snapshots' of the filter at two separate and *artificial* situations, neither of which may represent a 'physical' state along any point of the model.

It is possible to replace in the Levinson recursion the variables $A_n(z), A^{\#}_n(z)$ (forward and backward wave variables) by their sums and differences (the "pressure" and "volume velocity" variables):

$$A_n(z) + A^{\#}_n(z) \quad , \quad A_n(z) - A^{\#}_n(z), \qquad (2)$$

and use just one (or both) of them in alternative and more efficient forms of the Levinson algorithm. Since these two new variables may be regarded as representing the sound wave's pressure and volume-velocity, whose ratio forms an immittance function, they are called "immittance variables" [4].

The ISP parameters stem from the following function, which, in the pseudo model of the vocal tract, would represent the immittance at the glottis:

$$\mathcal{I}_p(z) = \frac{A_p(z) - A^{\#}_p(z)}{A_p(z) + A^{\#}_p(z)} \qquad (3)$$

The ISP parameters are the set of $p-1$ alternating locations on the unit circle $|z| = 1$ ('frequencies') of the (non-trivial) poles and zeros of $\mathcal{I}_p(z)$, and its 'gain' (the limit of $\mathcal{I}_p(z)$ as $z \to \infty$). For a stable $1/A_p(z)$, $\mathcal{I}_p(z)$ is thus featured by an ordered sequence of strictly increasing $p-1$ 'frequencies' and one positive gain $g = (1+k_p)/(1-k_p)$. This characterization represent conditions that are both necessary and sufficient for stability.

3. Adaptive Uniform Quantizer Design

If a uniform quantizer is used to code each parameter within the parameters vector, then each of the quantizers has to be adaptated to the number of bits that are allocated to this parameter. It is well known [6] that the quantizer's optimum scale relies upon both the number of bits and the probabilistic nature of the parameter to be quantized. Denote the variance of the parameter to be quantized by σ_x and the overload levels of the quantizer as $\pm x_{ol}$, then the *loading factor* is defined as $f_l := x_{ol}/\sigma_x$. A distortion function usually has a minimum with respect to f_l, which is a function of the parameter's probability density function and the number of allocated bits. Apparently, $f_{l,opt}$ is monothonically increasing with the allocated number of bits [6].

In general, it is impossible to calculate explicitly the optimum quantizer's step-size for a certain number of bits, unless numerical methods are being used. Furthermore, an allocation algorithm (e.g. the steepest descent algorithm [7, 9]) is influenced also by the *relative* distortion that each parameter introduces: more bits are allocated to more 'sensitive' parameters. The quantizer design is therefore a bi-dimentional minimization problem. When building the R-D curve, a simultanous distortion minimization search is necessary to maintain both the best allocation and optimum quantization scales, with the only a-priori knowledge that $f_{l,opt}$ may only increase with the allocated number of bits.

The algorithm that has been applied in our experiments uses a piecwise optimization, in which the allocation and the overload optimizations have been seperated. In each recursion step, a bit is allocated first using the steepest-descent algorithm. After a bit is allocated to a certain parameter, the overall overload factor f_l is raised, and the mean distortion is re-measured, to find whether the new quantization bounds yield higher performance. The process is repeated untill no further improvement is acheivable, after which the next bit is allocated and so on, untill all the requested number of bits are allocated. The step size of the bounds adaptation was $\Delta f_l = 0.1$, which has been found optimal for proper adaptation results. Given a certain model and quantization scheme, the bit allocation that results in 1 dB mean distortion and the overload point f_l (for two or three subgroups) are the only parameters that characterize the quantizer. This characterization is considerably less memory consumpting, compared to a non-uniform quantizer, in which all the quantization levels should be stored for each scalar quantizer.

Table 1: Experimental Conditions.

Data Base	approx. 4 minutes extracted from TIMIT
Contents	24 speakers from 8 USA regions 2 different sentences per speaker
Sampling	16 bit linear PCM, 16 KHz sampling decimated to 8 KHz
Pre-Processing	DC removal No pre-emphasis No silence removal
Framing	20 mSec Hamming windowed No overlapping
Analysis	10'*th* order LPC Burg's algorithm
Post Processing	No bandwidth expansion.
Distortion	Cepstral Distance

4. ISP vs. LSP quantization results

In order to compare the relative merits of ISP and LSP, several coding techniques have been implemented and tested using a speech data base. The experimental conditions were set as in [2], and are summarized in table 1.

In [2], an interframe-differentiation scheme has been reported to perform with an advantage of one bit per frame in favor of ISP. Additional subsequent experiments have been held [5], to assess the value of the ISP parameters (compared to LSP) in also other coding schemes, as follows:

1) Direct Uniform Quantization (UQ).

Table 2: Spectral Distortion Statistics.

Coder		ISP	LSP
Direct	**bits @ 1 db**	37	38
	3 db outliers	1.41 %	1.64 %
	5 db outliers	0.25 %	0.39 %
Intra Frame	**bits @ 1 db**	35	36
	3 db outliers	0.73 %	1.13 %
	5 db outliers	0.03 %	0.04 %
AQFW	**bits @ 1 db**	35	36
	3 db outliers	1.13 %	1.18 %
	5 db outliers	0.18 %	0.20 %
NUQ	**bits @ 1 db**	35	36
	3 db outliers	0.40 %	0.53 %
	5 db outliers	0.01 %	0.09 %
DCT-Inter.	**bits @ 1 db**	32	33
	3 db outliers	2.37 %	2.58 %
	5 db outliers	0.65 %	0.85 %

2) Intra-frame differentiation, following Soong & Juang's dLSP in [7].

3) Forward Sequential Adaptive Quantization, following Sugamura & Farvardin's AQFW in [8].

4) Non-Uniform Quantization (NUQ).

5) Transform Inter-frame Quantization. This scheme is after Farvardin & Laroia's Hybrid DCT-PCM coder in [9], except that we use uniform quantization.

The rate-distortion curve for each of the above schemes are depicted in the enclosed figures 1-5. Table 2 summarizes additional details, regarding distortion conditions for the bit rate at which the mean distortion level does not exceed 1 dB. The 1 dB distortion level with reasonably low outlier rates is an often accepted criteria for "transparent" quantized speech quality [10]. It is seen that ISP outperforms LSP by one bit in all the examined coding experiments. Moreover, for the ISP quantizer, the achievement of 1 dB distorsion level with 1 bit less is always accompanied by also less outliers percentage than the LSP at its higher bit rated 1 dB distorsion level.

5. References

[1] F. Itakura "Line spectrum representation of linear predictive coefficients" *J. Acoust. Soc. AM.* 535(a), s35(A), 1975

[2] Y. Bistritz and S. Peller, "Immittance Spectral Pairs (ISP) for Speech Encoding," *Proc. Int. Conf. on Acoustics, Speech and Signal Proc.*, pp. II-9 - II-12, April 1993.

[3] Y. Bistritz, "Immittance Domain Modeling for Linear Predictive Coding," in preparation.

[4] Y. Bistritz, H. Lev-Ari and T. Kailath, "Immittance-Domain Levinson Algorithms", *IEEE Trans. Inform. Theory.*, vol. IT-35, May 1989. See also *Proc. Int. Conf. on Acoustics, Speech and Signal Proc.* pp. 253-256, Tokyo, Japan, Apr. 1986.

[5] S. Peller, "Immittance Spectral Models for Speech Encoding," M.Sc. Dissertation, Tel-Aviv University, 1993.

[6] N. S. Jayant and P. Noll, *"Digital Coding of Waveforms, Principles and Applications to Speech and Video"*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1984.

[7] F. K. Soong and B-H Juang, "Optimal quantization of LSP parameters" *IEEE Trans. Speech and Audio Process.* SAP-1, pp. 15-24, 1993. Also in *Proc. Int. Conf. on Acoustics, Speech and Signal Proc.*, pp. 394-397, 1988.

[8] N. Sugamura and N. Farvardin, "Quantizer Design in LSP Speech Analysis and Synthesis," *Proc. Int. Conf. on Acoustics, Speech and Signal Proc.*, pp. 398-401, 1988.

[9] N. Farvardin and R. Laroia "Efficient encoding of speech LSP parameters using the discrete cosine transformation" *Proc. Int. Conf. on Acoustics, Speech and Signal Proc.* pp. 168-171, 1989.

[10] K. K. Paliwal and B. S. Atal "Efficient vector quantization of LPC parameters at 24 bits/frame" in *IEEE Trans. Speech and Audio Process.*, SAP-1 , pp. 3-14, 1993. Also in *Proc. Int. Conf. on Acoustics, Speech and Signal Proc.* pp. 661-664, 1991.