Interpolation of Long Gaps in Audio Signals Using Line Spectrum Pair Polynomials

Paulo A. A. Esquef
HUT – Laboratory of Acoustics and Audio Signal Processing
P.O.Box 3000, FIN-02015 HUT, Espoo, Finland

paulo.esquef@hut.fi

Abstract

This technical report addresses model-based interpolation of long signal gaps. It demonstrates that employing a modified autoregressive AR model, computed as a weighted sum of line spectral pair (LSP) polynomials, is more efficient computationally than using a conventional AR model, since longer signal gaps can be interpolated at reduced model order.

Key-words: acoustic signal processing, audio reconstruction, AR models, line spectrum pair polynomials

1. INTRODUCTION

Reconstruction of missing samples of audio and speech signals is needed in many real-life digital signal processing applications. Signal losses can be caused by several factors and reasons. For instance, transmission errors through digital channels can lead to signal drop-outs due to packet losses [1]. Signal losses may occur already in the analog domain prior to the signal digitalization. These are mainly caused by damages on the medium that stores the analog signal. Typical examples are the clicks, pops, and crackles associated with old disk records [2].

Autoregressive (AR) modeling of audio and speech signals has been utilized in several applications, such as linear predictive coding [3], outlier detection [4], and signal reconstruction [5, 6]. AR-based interpolation methods are usually suitable for interpolating relatively short gaps. This restriction comes from the fact that, in general, musical signals are non-stationarity. Indeed, stationary can only be accepted as a valid hypothesis within frames of about 20 ms. Thus, the signal has to be segmented in short frames to which separate AR models are fitted. If a single AR model is to be fitted to a segment with missing samples, the remaining number of known samples in the segment should be enough to guarantee the statistical significance of the estimated model. Of course, the larger the number of missing samples in the segment, the poorer the estimated model and the performance of the reconstruction algorithm.

Furthermore, AR-based interpolation of long gaps shows poor performance toward the middle of the gap [5]. This is due to the LS minimization of the modeling error, used in the solution for the unknown samples. This criterion tends to yield a model excitation whose variance is too low. In the ideal case of a null excitation, the resulting interpolated signal would be mainly governed by the impulse response of the AR model. For a stable AR model, the closer its poles are to the unit circle, the faster its impulse response decays over time. Thus, if the gap to be concealed lasts much longer than the time over which the impulse response of the AR model conveys most of its energy, there will be little signal energy toward the middle of the gap. Increasing the model order tends to push the poles of the resulting AR model closer to the unit circle. Hence, this resource helps to alleviate the energy vanishing effect, since an AR model with a slower decaying impulse response is achieved. Other solutions to this problem consist of imposing a limit to the minimization of the prediction error variance [7, 8].

When the gap to interpolate becomes too long it may happen that the spectral characteristics of the signal before and after the gap differ substantially. In these cases, a more suitable solution consists of employing two separate AR models, one for the fragment preceding the gap and another for the samples that succeed the gap. Such a scheme has been proposed in [9], which presents a weighted LS solution for the missing samples given the two separate AR models. The same rationale of employing two different AR models for interpolation of long gaps has been presented in [10, 11, 12].

Interpolation of long signal fragments has also been proposed through sinusoidal modeling [13]. In this case, the gap information can be inferred from the parameters of the sinusoidal tracks that surround the gap. The method has been reported effective for concealing gaps of up to 30 ms, or even longer, depending on the signal stationarity [13].

In this work, interpolation will be carried out through relatively low-order AR-based modeling. However, a modified AR model will be used in the interpolation method as replacement to the conventionally estimated AR model. The modification consists of first estimating each conventional AR model and obtaining its associated LSP polynomials [14]. Then, a modified AR model is computed from a weighted sum of the LSP polynomials [14]. The advantage of this solution is that a single weight factor can be adjusted to place the poles of the modified model closer (or on) the unit circle, regardless of the order of the original AR model. Thus, the impulse response of the modified model will decay slower than that of the conventional model of same order, offering an immediate benefit to the interpolation problem.

The side-effect of using the weighted LSP approach is that the resonance frequencies of resulting AR models become biased, especially in the high-frequency range. As a result, some distortions are perceived within the reconstructed portions of signal. Nevertheless, the modified models can produce a perceptually better interpolation result than the one attained via conventional AR model of the same order. The gain comes from the preservation of the signal energy across the gap, which is perceptually important. The relatively low-order needed for the LSP-modified AR model to carry out the interpolation task accounts for a significant reduction in the computational complexity of the interpolation method.

This paper is organized as follows. Section 2 defines the line spectral pair polynomials as well as reviews their properties. Section 3 reviews the concept and properties of the modified AR model, which is based on the weighted sum of LSP polynomials. Section 4 describes the interpolation method used in the simulations. Experimental results are shown in Section 5. Performance comparisons among the results of the interpolation method when using the modified and the conventional AR models are provided in Section 6. Conclusions are drawn in Section 7.

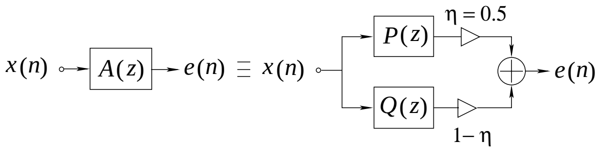

Figure 1: Equivalence between $A(z)$ and the LSP polynomials.

2. LINE SPECTRAL PAIRS POLYNOMIALS

Let the response of a p^{th}-order prediction filter be defined as $A(z) = 1 - \sum_{k=1}^{p} a(k)z^{-k}$, where $a(\cdot)$ is the direct form of the prediction coefficients. The transfer function of the synthesis filter is $1/A(z)$, which is guaranteed to be stable if $A(z)$ is of minimum phase.

The idea behind LSP polynomials is to map the polynomial $A(z)$ into two other equivalent polynomials, whose zeros are located on the unit circle. The mapping is given by $P(z) = A(z)+z^{-(p+1)}A(z^{-1})$ and $Q(z) = A(z)-z^{-(p+1)}A(z^{-1})$, where the polynomials $P(z)$ and $Q(z)$ are called LSP polynomials [15]. Moreover, $A(z)$ can be reconstructed from the LSP polynomials via $A(z) = \frac{1}{2}\left[P(z) + Q(z)\right]$. Figure 1 illustrates the equivalence between $A(z)$ and the weighted sum of $P(z)$ and $Q(z)$.

3. WEIGHTED SUM OF LSP POLYNOMIALS

A linear prediction scheme based on a weighted sum of the LSP polynomials has been proposed in [14]. The prediction filter is defined as

$$D(z, \eta) = \eta P(z) + (1 - \eta)Q(z), \tag{1}$$

where, as before, $P(z)$ and $Q(z)$ are the symmetric and antisymmetric LSP polynomials associated with a p^{th}-order predictor $A(z)$. The weight η is a real-valued gain. As seen in Section 2 and in Figure 1, for $\eta = 0.5$ the equivalence $D(z, \eta = 0.5) = A(z)$ holds true. It should be noted that for $\eta \neq 0.5$, $D(z, \eta)$ has order $p + 1$.

For $\eta = 1$ and $\eta = 0$ the modified predictor $D(z, \eta)$ reduces to the symmetric and antisymmetric LSP polynomials, respectively. Therefore, for these values of η the poles of $1/D(z, \eta)$ are on the unit circle. Moreover, it is shown in [14] that $1/D(z, \eta)$ is guaranteed to be stable if $\eta \in \,]\,0, 1\,[$. The AR model $1/D(z, \eta)$ will be referred hereafter to as the WLSP-modified AR model.

Figure 2 shows the root *locus* of $1/D(z, \eta)$ as function of η for a synthetic 4^{th}-order AR model. Visual inspection on the root *locus* reveals that setting η close to 1 results in an all-pole model $1/D(z, \eta)$ whose resonance frequencies (in the low-frequency range) are closer to those of the original model $1/A(z)$ than if η is set close to 0. Besides, setting $\eta = 0$ implies a pole at DC, which may be problematic for synthesis applications, since there is always some DC level in the excitation signals. Conversely, setting $\eta = 1$ implies a pole at $z = -1$. This pole, however, seems to be less harmful perceptually than that at $z = 0$.

It will be seen in Section 5 that pushing the poles of the AR models used in the interpolation method close to the unit circle, possibly without major changes in their resonance frequencies, favors the interpolation performance. This can be done via a proper choice of the η parameter in $1/D(z, \eta)$. Alternatively, one could think of solving for the roots of $A(z)$, artificially increasing the radius of the poles, and then recomputing the model coefficients. However, solving for the roots of high-order polynomials is a not only prone

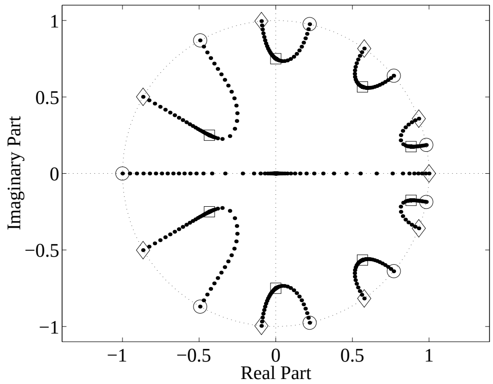

Figure 2: Root *locus* of $1/D(z, \eta)$ as function of $0 \leq \eta \leq 1$ for a synthetic 4th-order AR model. The markers □, ◊, and ◯ indicate the pole locations for $\eta = 0.5$, $\eta = 0$, and $\eta = 1$, respectively.

to numerical errors, but also a computationally demanding task. Conversely, the computational burden to obtain the WLSP-modified models is negligible: it requires only p multiplications and additions per model.

4. MODIFIED INTERPOLATION METHOD

The WLSP-modified AR models can be just plugged into any AR-based interpolation method, e.g., the least-squares AR (LSAR) method described in [6]. As the task involved here is interpolation of long gaps, one can profit from using two different AR models, one for the fragment that precedes the gap and another for the segment that succeeds the gap [9, 11]. The AR-based interpolation scheme described in [11] will be used in this work. Supposing a gap of G samples in between two segments of N samples each, the steps of the interpolation method follow.

1. Estimate an AR model for the N-sample-long segment that immediately precedes the gap and compute the WLSP-modified AR model from it using η close to 1.

2. Forward extrapolate the signal across the gap by exciting the WLSP-modified AR model, as described in [12]. The procedure can be equivalently accomplished by a) computing an excitation via inverse filtering the segment upon which the model was estimated through the WLSP-modified AR model; b) appending G zeros to the end of the resulting excitation; and c) exciting the modified AR model with this zero-padded excitation. The last G output samples of the previous filtering correspond to the desired extrapolated signal.

3. Repeat the previous two items for the segment that succeeds the gap. This time, however, a backward signal extrapolation is aimed for. Hence, appropriate time-reversals of the signals involved should be performed.

4. Cross-fade the two extrapolated sequences via the generalized window[1] as described in [11].

[1]The roll-off parameter is set to 3 in all simulations shown.

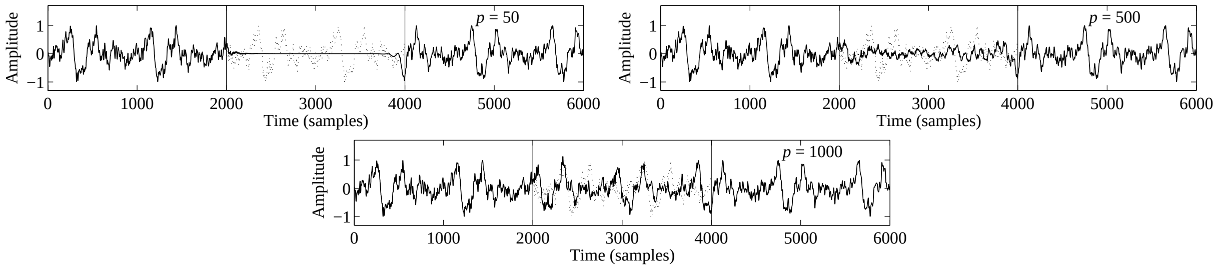

Figure 3: Performance of the interpolation method for several model orders. The gap is located between samples 2001 and 4000. The original signal within the gap is shown in dotted line.

5. EXPERIMENTAL RESULTS

The purpose of this section is to compare visually the performance of the interpolation method under different model setups. The results attained when using conventional AR models will be confronted against those attained when using WLSP-modified AR models. The idea is to first illustrate the vanishing energy effect in interpolation of long signal gaps as well as the role of the model order in this matter. Then, the benefit of using WLSP-modified models is demonstrated. Of course, purely visual inspections do not tell much about the sound quality the results. This issue is covered in Section 6.

5.1. Case Study: Low-frequency piano tone

This case study features a low-pitch piano tone (fundamental frequency $f_0 \approx 50$ Hz, played *forte*). Such a tone has hundreds of partials, and thus, modeling it as an AR process requires high-order models. Applying an insufficient model order not only leaves some mode resonances unaccounted for, but also tends to enlarge the bandwidth of the modeled ones, i.e., the poles associated with these modes are not as close to the unit circle as they should be.

The pole locations of a stable AR model define how fast its impulse response decays over time. In its turn, the decay profile of the the model response has a crucial impact on the performance of model-based interpolation of long gaps. For the extrapolation-based reconstruction scheme used here, this issue is even more relevant, since the extrapolated signals are basically the unforced response of the AR models, for given initial conditions. Thus, fast decaying impulse responses may not suffice to fill in a long gap completely. The result is then a reconstructed signal whose energy fades out towards the middle of the gap.

First, the effect of the model order on the interpolation performance is investigated. A gap of 2000 samples is artificially created in the signal, starting 5000 samples after the note onset. Figure 3 compares the interpolation results for several choices of model order. The AR models are estimated using Burg's method [16]. The set of 2000 samples that precedes the gap is used to estimate the forward AR predictor. Similarly, the set of 2000 samples that succeed the gap is used to estimate the backward AR predictor. From Fig. 3 it is clear that model orders below 500 are insufficient to properly reconstruct this particular signal. However, setting $p = 1000$ seems sufficient for most cases [10].

Now, the conventional AR models are replaced with their WLSP-modified versions.

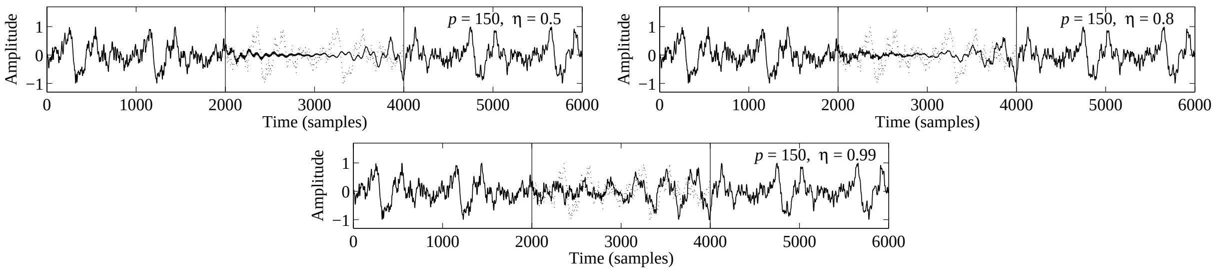

Figure 4: Performance of the proposed interpolation method for model order $p = 150$ and several values of η. The gap is located between samples 2001 and 4000 and the original signal within the gap is shown in dotted line.

The model order[2] is kept at $p = 150$ and $\eta \in \{0.5, 0.8, 0.99\}$. The interpolation results are summarized in Fig. 4. It can be observed that setting η close to 1 seems to improve the interpolation performance, in that the energy of the interpolated sequences tends to be preserved. In this regard, the result of interpolation method, when employing the WLSP-modified models ($\eta = 0.99$) of order $p = 150$, is similar to that achieved when using the conventional AR models ($\eta = 0.5$) of order $p = 500$.

The previous result counts favorably to the usage of the WLSP-modified models in the interpolation scheme, in terms of a better balance between interpolation performance and computational efficiency. Indeed, the extra burden needed to compute the LSP polynomials and the WLSP-modified models, is negligible compared to the costs involved in the model estimations and the extrapolation procedures.

The side-effect of using the WLSP-modified models in the interpolation scheme is that they produce interpolated signals whose resonance frequencies may be biased compared to those of the original signal. Nonetheless, slightly frequency-biased mode resonances are likely to be initially excited with an energy level as high as that of the original resonances. This is why the interpolation method still works. Note that strongly biased resonances will probably be excited at low energy levels. Hence, regarding contributions to audible distortions in the reconstructed signals, strongly biased resonances may be less harmful than slightly biased ones.

Figure 5 confronts the spectrum of the original signal within the gap against the spectra of the reconstructed signals. One verifies that the original signal has a large number of spectral peaks. Moreover, the spectrum of the interpolated signal via $p = 150$ and $\eta = 0.5$ reveals that only the most prominent spectral peaks are modeled. These peaks are not as sharp as those of the original spectrum.

On the other hand, the modeling setup $p = 150$ and $\eta = 0.99$ offers an interpolated signal with a larger number of prominent peaks. The peaks are also sharper than those of the signal attained via the modeling setup $p = 150$ and $\eta = 0.5$. However, it is evident that the complex spectral structure of the original signal is not well modeled by any of the previous configurations that employ low-order models.

As for the frequency of the resonance modes, the spectral peaks of the interpolated signal, attained via the modeling setup $p = 150$ and $\eta = 0.99$, show a good correlation to the those of the original signal, at least visually. Figure 6 illustrates better this fact. Nevertheless, frequency deviation in some spectral peaks can be observed. For instance, the frequency of the prominent peak that appears around 1500 Hz in the original signal is higher in the model spectrum. However, the low-frequency peaks seem to be less biased.

[2]The value of p refers to the order of the original AR model. The order of the WLSP models is $p + 1$.

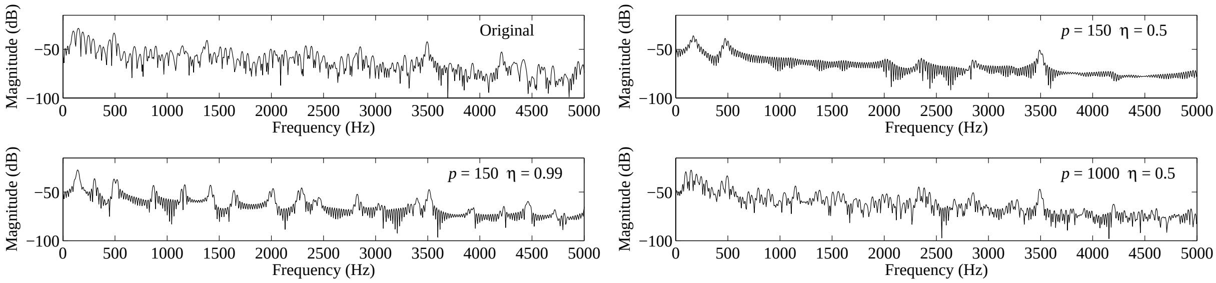

Figure 5: Detail (up to 5 kHz) of the spectra of the original and reconstructed signals.

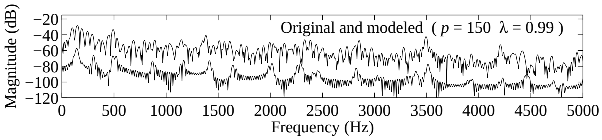

Figure 6: Detail (up to 5 kHz) of the spectra of the original and reconstructed signal under setup $\eta = 0.99$ and $p = 150$. The modeled spectrum is plotted with a shift of -30 dB for clarity.

Therefore, it is plausible to speculate that the quality of this interpolated signal can be acceptable, from the perceptual point of view, in spite of possible audible distortions.

6. PERFORMANCE ASSESSMENT

In real-life situations, performance assessment of restored audio can only be carried out via listening tests. As usually a referential uncorrupted version of the signal is absent, objective measures are prevented. Nevertheless, the source of degradation can be artificially emulated and used to corrupt clean versions of test signals. Thus, the quality of the restored signals can be evaluated objectively. Standard measures, such as signal-to-noise ratio, may fail to reflect the subjective quality of the results. An alternative is to employ objective measures based on psychoacoustics.

6.1. Perceptual Audio Quality Measurement

The Perceptual Audio Quality Measure (PAQM) [17] has been originally devised to assess the quality of lossy-coded audio signals. It has been demonstrated to correlate strongly with human judgment for both wideband audio and band-limited speech signals. The basic principle behind the PAQM is to compare the inner ear representations of the processed and reference signals. Based on these representations a cognitive model generates an index of overall dissimilarity between the two signals. This index, the PAQM, reflects both loss of valuable information as well as introduction of spurious artifacts in the processed signal.

Note that, the PAQM values shown in the simulations reported here are merely those of the dissimilarity index, and not subjective gradings of the results, such as the mean opinion score (MOS). Mapping the dissimilarity index into the MOS grading can be carried out [17]. However, such mapping would be not only application dependent, but also would require extensive listening tests to assure a high correlation between the PAQM and MOS values.

In absolute and objective terms, the lower the value of PAQM, the more similar the processed signal is to the reference. A $\mathrm{PAQM} = 0$ implies a processed signal perceptually identical to the reference. However, drawing conclusions on the subjective quality of a reconstructed signal based solely on its PAQM is not straightforward. In other words, the question arises: how close to zero should the PAQM be so that the restored signal can be considered perceptually indistinguishable from the original? At least, two references can be employed to help assessing the quality of the results in a comparative way: the PAQM of the corrupted signal (with gaps); and the PAQM of a satisfactorily restored signal, e.g., by adopting high-order ($p = 1000$) AR models within the conventional interpolation method. The PAQM of this signal gives a more realistic reference (than perfection) for a satisfactory interpolation performance. The PAQM of the corrupted signal allows verifying whether the interpolation is contributing or not to improve the perceptual quality of the interpolated signals.

The absence of a mapping into a subjective quality grading system gives the PAQM a speculative role in regard the subjective assessment of the results. In other words, no one knows how apart from each other two values of PAQM should be in order to the corresponding signals be perceived as perceptually different. Nevertheless, it is plausible to speculate that PAQM values close to each other are likely to reflect signals with similar perceived quality. In this context, the external referential PAQMs, i.e., those of the degraded and a satisfactorily restored signals, play an essential role in allowing an approximate inference on the quality of other restored versions of the signal from their PAQMs.

6.2. Test Signals

Finding a set of test signals that can represent a broad class of audio and speech signals is a difficult task. For convenience, in this work such set will be limited to four signals only. They are **Piano Tone**: a piano tone ($f_0 \approx 50Hz$) played *forte*; **Pop**: a 14-s-long excerpt of Finnish pop music; **Classic**: a 13-s-long excerpt of orchestral music; and **Singing**: a 20-s-long excerpt of pop singing *a capella*.

6.3. Experimental Setup

The experiments consist of artificially creating gaps in the test signals and reconstructing the signal in them. After that the PAQM related to each pair of original/reconstructed signal is measured. This way, one can assess the effects of the processing parameters on the interpolation performance.

The the length of the gaps is set to 2000 samples (@ 44.1 kHz), corresponding thus to duration of about 45 ms. This poses a challenge for the reconstruction procedure, since such length is on limit above which the signal frame can no longer be considered stationary.

Creating gaps in a periodical fashion allows evaluating the interpolation performance in portions of the same signal with different temporal and spectral characteristics. Moreover, the periodical mutes are quite annoying perceptually. As a measure based on psychoacoustics, the PAQM is likely to evaluate the ability of the interpolation method to restore or not the sense of continuity in the interpolated signals.

The idea of the experiments is to check whether the usage of WLSP-based models leads to any improvement in the performance of the interpolation method, i.e, a better balance tradeoff between perceptual quality or computational efficiency. Rather than evaluating

Table 1: Parameters of the experimental setup.

Parameter	Value
Gap length	$G = 2000$ samples
Gap periodicity	50000 samples
Order (WLSP models)	$p = 150$
Order (conv. AR models)	$p \in \{150, 300, 500, 1000\}$
AR estimation	Burg's method, $N = 2000$ samples
AR modification	WLSPs with $\eta \in [0.5, 1.0]$

the results for various model orders, the investigation is focused on the effect of η on the quality of the restored signals. The order of the WLSP-based models is then arbitrarily set to 150.

For the reasons given in Section 6.1, the PAQMs of the test signals restored via the conventional interpolation method with $p \in \{150, 300, 500, 1000\}$ are provided as a means to aid performance comparisons. The processing parameters used in the experimental setup are summarized in Table 1.

6.4. Effect of η

The performance assessment is carried out using the experimental setup defined in the previous section. As seen in Fig. 4, the effect of η on the interpolation results is more dramatic when $\eta \to 1$. Therefore, when measuring the PAQM of the results as function of η, it is reasonable to adopt a non-uniform distribution for the values of η within the range between 0.5 and 1. In this case, a higher sample density should be employed as η approaches 1. The values of η were arbitrarily chosen as $\eta \in \{0.5,\ 0.8,\ 0.9,\ 0.99,\ 0.999,\ 0.9999,\ 1\}$.

The solid lines in Fig. 7 show the PAQMs of the restored signals as function of η. In all cases, the model order was fixed to $p = 150$. The horizontal dashed lines indicate the values of PAQM attained using conventional AR models the indicated orders. As additional references, the PAQMs related to the corrupted version of the signals are also shown.

From Fig. 7 one observes that, at least for the test signals employed, the PAQM decays monotonically as the value of η varies from 0.5 to 1. From the perceptual point of view, the best results in all examples are attained when $\eta = 1$. The decrease in perceptual quality due to the presence of the gaps seems to be similar for all test signals. As one could anticipate, increasing the order of the conventional models improves the quality of the results. For the signals **Piano Tone** and **Classic**, signal reconstruction with conventional AR models of order $p = 150$ already yields some improvement in the perceptual quality. This improvement is less substantial for the signal **Singing** and even null for the signal **Pop**. In this example, the restored version exhibiting a slightly higher PAQM than the degraded signal does not necessarily mean that a further degradation has taken place. Recalling the remarks provided in Section 6.1, both signals should be regarded as perceptually similar.

According to the PAQM, for the signal **Piano Tone**, the interpolation result attained when employing WLSP-modified AR models ($\eta = 1$) of order $p = 150$ is comparable to that obtained through conventional AR models of order $p = 500$. The PAQMs show, however, that a better interpolation performance for this signal can be achieved by employing higher-order conventional AR models. Similar conclusions can be drawn from the PAQM

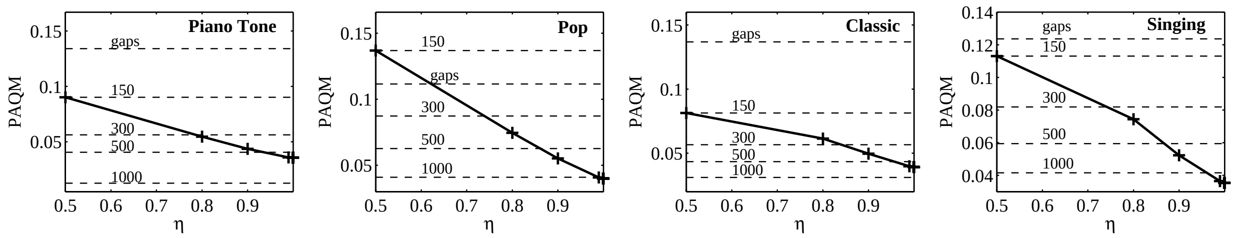

Figure 7: PAQMs of the interpolated signals under several parameter setups. The solid lines represent the PAQM as function of η with $p = 150$. The dashed lines refer to the PAQM of the results attained with $\eta = 0.5$ and the indicated model orders. The PAQM of the corrupted signal is also shown (gaps).

results associated with the restored versions of the signal **Classic**.

Interestingly, for the signals **Pop** and **Singing**, the performance of the interpolation method when using WLSP-modified AR models ($\eta = 1$) of order $p = 150$ surpasses that attained through conventional AR models of order $p = 1000$. In all cases, using the WLSP-modified AR models allows adopting significant lower model orders than those needed if conventional AR models were used to achieve similar interpolation performances. Therefore, the modified models substantially increase the computational efficiency of the interpolation method.

6.5. Subjective Evaluation

Informal listening tests on the restored signals have been conducted to confront the PAQM in subjective terms. The results reveal that, despite some audible distortions, the reconstructed signals attained by the WLSP-based approach with $p = 150$ and $\eta = 1$ sound better than those by the conventional method with $p = 150$. However, the overall subjective impression was that the quality of the results achieved by the conventional interpolation with $p = 1000$ was always, perceptually, slightly higher than that of the signals restored with the modified scheme with $p = 150$ and $\eta = 1$.

The latter finding is in agreement with the PAQMs computed for the corresponding versions of the signals **Piano Tone** and **Classic**. However, for the signals **Pop** and **Singing**, the related PAQMs contradict the results of the subjective evaluation. That discrepancy may be explained by the approximate character of the PAQM comparisons, as used here, to draw conclusions on the subjective quality of the results. Moreover, the accuracy of the PAQM to reflect perceptual quality tends to decrease as the dissimilarity between the processed signal and the reference increases. In this regard, one should notice that signals **Pop** and **Singing** are more non-stationary in nature than signals **Piano Tone** and **Classic**. Therefore, adequate signal reconstruction of the former signals is a more challenging task. Consequently, the satisfactorily restored versions ($p = 1000$) of signals **Pop** and **Singing** are not as close to the reference as those of the signals **Piano Tone** and **Classic**. Hence, for the former signals, less confident values of PAQM are expected to occur. Sound examples are available at on the web URL: `http://www.acoustics.hut.fi/publications/papers/tassp-int/`

7. CONCLUSIONS

This work presented an application of LSP polynomials to signal interpolation across long signal gaps. In particular, the conventional AR models used in the interpolation method were replaced with modified AR models, which are based on a weighted sum of LSP polynomials. The pole location of these modified models can be controlled through the weighting parameter.

Experiments were conducted on a set of four test signals. Objective measures taken over the restored signals revealed that adopting a low-order AR model and setting the weight parameter to 1 yield better interpolation results. This choice corresponds to replace the conventional AR model with its symmetric LSP polynomial, which has all its poles on the unit circle.

Subjective evaluations via informal listening tests showed that the proposed modification introduces some distortions in the interpolated signals and their perceptual quality is inferior to that of the satisfactorily restored signals. Nevertheless, for low-order models, the modified scheme leads to a substantial improvement in perceptual quality compared to the conventional method with equal order. Since the proposed modification implies only a negligible extra computational cost, a better balance between interpolation performance and computational efficiency can be claimed.

Acknowledgments

The work of Paulo A. A. Esquef has been supported by the Brazilian National Council for Scientific and Technological Development (CNPq-Brazil). The author wishes to thank Mr. Tom Bäckström and Dr. Luiz W. P. Biscainho, for providing some of the codes used in the simulations, as well as Dr. Vesa Välimäki for proofreading the manuscript.

8. BIBLIOGRAPHY

[1] N. S. Jayant and S. Christensen, “Effects of Packet Losses in Waveform Coded Speech and Improvements Due to an Odd-Even Sample-Interpolation Procedure,” *IEEE Trans. Communications*, vol. CAM-29, pp. 101–109, Feb. 1981.

[2] P. Wilson, “Record Contamination: Causes and Cure,” *J. Audio Eng. Soc.*, vol. 13, pp. 166–176, Apr. 1965.

[3] J. Makhoul, “Linear Prediction: A Tutorial Review,” *Proc. IEEE*, vol. 63, no. 4, pp. 561–580, 1975.

[4] S. V. Vaseghi and P. J. W. Rayner, “Detection and Suppression of Impulsive Noise in Speech Communication Systems,” *IEE Proceedings*, vol. 137, pp. 38–46, Feb. 1990.

[5] A. J. E. M. Janssen, R. N. J. Veldhuis, and L. B. Vries, “Adaptive Interpolation of Discrete-Time Signals That Can Be Modeled as Autoregressive Processes,” *IEEE Trans. Signal Processing*, vol. ASSP-34, pp. 317–330, Apr. 1986.

[6] S. J. Godsill and P. J. W. Rayner, *Digital Audio Restoration — A Statistical Model Based Approach*, ch. 5. London, UK: Springer-Verlag, 1998.

[7] P. J. W. Rayner and S. J. Godsill, "The Detection and Correction of Artefacts in Degraded Gramophone Recordings," in *Proc. IEEE ASSP Workshop Applications Signal Processing Audio Acoustics*, pp. 151–152, Oct. 1991.

[8] M. Niedźwiecki, "Statistical Reconstruction of Multivariate Time Series," *IEEE Trans. Signal Processing*, vol. 41, pp. 451–457, Jan. 1993.

[9] W. Etter, "Restoration of a Discrete-Time Signal Segment by Interpolation Based on the Left-Sided and Right-Sided Autoregressive Parameters," *IEEE Trans. Signal Processing*, vol. 44, pp. 1124–1135, May 1996.

[10] I. Kauppinen, J. Kauppinen, and P. Saarinen, "A Method for Long Extrapolation of Audio Signals," *J. Audio Eng. Soc.*, vol. 49, pp. 1167–1180, Dec. 2001.

[11] I. Kauppinen and J. Kauppinen, "Reconstruction Method for Missing or Damaged Long Portions in Audio Signal," *J. Audio Eng. Soc.*, vol. 50, pp. 594–602, July/Aug. 2002.

[12] I. Kauppinen and K. Roth, "Audio Signal Extrapolation – Theory and Applications," in *Proc. 5th Int. Conf. on Digital Audio Effects*, (Hamburg, Germany), pp. 105–110, Sept. 2002. URL: http://www.unibw-hamburg.de/EWEB/ANT/dafx2002/papers.html.

[13] R. C. Maher, "A Method for Extrapolation of Missing Digital Audio Data," *J. Audio Eng. Soc.*, vol. 42, pp. 350–357, May 1994.

[14] P. Alku and T. Bäckström, "All-pole Modeling Technique based on the Weighted Sum of the LSP Polynomials," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2002)*, vol. 1, (Orlando, Florida, USA), pp. 665–668, May 2002.

[15] F. K. Soong and B. W. Juang, "Line Spectrum Pair (LSP) and Speech Data Compression," in *Proc. Int. Conf. Acoust., Speech, Audio Signal Procesing*, (San Diego, CA), pp. 1.10.1–1.10.4, 1984.

[16] M. H. Hayes, *Statistical Signal Processing and Modeling*, ch. 6. John Wiley & Sons, Inc., 1996.

[17] J. G. Beerends, "Audio Quality Determination Based on Perceptual Measurements Techniques," in *Applications of Digital Signal Processing to Audio and Acoustics* (M. Kahrs and K. Brandenburg, eds.), ch. 1, Kluwer Academic Publishers, 1998.