AN ALGORITHM FOR WIDEBAND AUDIO CODING BASED ON LMS/RLS SWITCHED PREDICTORS

Enzo Mumolo *Mauro Scagnol* *Alberto Carini*
DEEI, Dipartimento di Elettrotecnica, Elettronica ed Informatica
Universita' di Trieste, via Valerio 10, 34127 Trieste, Italy
Tel: +39 40 6763861; fax: +39 40 6763460
e-mail: mumolo@univ.trieste.it

ABSTRACT

This paper describes an algorithm for coding high quality audio signals using a switched ADPCM approach. Several theoretical and practical issues are considered as outlined in the paper.

1 INTRODUCTION

Nowadays, there is a rapidly growing number of commercial application which require exchange of audio information. One of the approaches for coding audio signals is based on ADPCM algorithms. Besides the ISO/MPEG standard [1], there are currently several activities on ADPCM-based approaches, such as the Intel/DVI, the Microsoft Wav-ADPCM or the ODA standards - for embedding audio signals into multimedia documents.

The basic idea of the described algorithm is to use RLS based adaptive prediction instead of LMS in ADPCM coders. As a matter of fact, the performance improvement obtained with RLS is very high. Of course, RLS introduces a number of problems which has been solved with a switching mechanism, as summarized below.

The use of switched predictors in signal coding is an old idea. Differently from early applications, however, we used a switched scheme for facing the stability problems which arise when RLS is used in an ADPCM framework.

One fundamental characteristic of RLS approches is that they are not influenced by the statistical properties of the signal as in LMS. Apart particular cases [8], generally the convergence speed of RLS is much higher than that of LMS. For this reason, ADPCM algorithms with RLS adaptation can be used with high bandwidth signals, such as audio signals. The tracking capabilities of RLS is the reason why there are stability problems in the coding system, as it will be explained more formally in the paper.

Finally, it will be shown that the performance of ADPCM systems increase with frequency; if the predictor is able to track the fast variations of the signal, the performance of ADPCM would be very high.

The switching mechanism leads to an algorithm for audio coding with the following main features:

- It has no coding delay;
- It has an $O(N)$ complexity, N being the predictors order;
- The algorithm's basic scheme leads to performance improvement over the G.722 coding algorithm;
- The switched mechanism yields a variable bit rate coding system. Since, if the bit rate is lowered, the performance degradation is quite smooth, the algorithm can be used in embedded coding applications;
- It is a starting point for improved schemes which might be obtained by using noise-shaping approaches.

The paper is organized as follows. In Section 2, the algorithm is described in greater details and some theoretical results are reported. In Section 3 we report some experimental results and in Section 4 some concluding remarks are discussed.

2 The coding algorithm

The block diagram of the algorithm is depicted in Fig.1. It is worth noticing that there are two different stability issues in this system. One is concerned with the numerical stability of the RLS algorithm, and the other is concerned with the input-dependent stability of the coding system. The switching mechanism is aimed at solving the latter type of instability.

Remark 1 In order to avoid possible mistracking problems, the updating step of the quantizers is shared between the two subsystems.

2.1 RLS Algorithms

Though characterized by a fast tracking behaviour, RLS algorithms have numerical stability problems. Many approaches to overcome that problem have been derived, such as the FTF, the Fast Lattice and the QR-type algorithms [7]. In this work, however, we used a Square-Root types RLS algorithm.

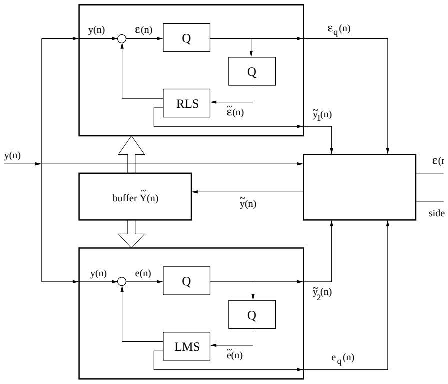

Figure 1: Block diagram of the algorithm

To start dealing with them, let us first recall that RLS-SQR algorithm is computed by the relations described in Table I [6]:

eq1	$F(n) = S^T(n-1)X(n)$
eq2	$L(n) = S(n-1)F(n)$
eq3	$\beta(n) = \lambda + F^T(n)F(n)$
eq4	$\alpha(n) = 1/[\beta(n) + (\lambda\beta(n))^{1/2}]$
eq5	$S(n) = 1/\sqrt{\lambda}[S(n-1) - \alpha(n)L(n)F(n)^T]$
eq6	$\epsilon(n) = y(n) - H^T(n-1)X(n)$
eq7	$\epsilon_q(n) = Q^{-1}[Q[\epsilon(n)]]$
eq8	$\tilde{\epsilon}(n) = \epsilon_q(n)/\beta(n)$
eq9	$H(n) = H(n-1) + \tilde{\epsilon}(n)L(n)$
eq10	$\tilde{y}(n) = \lambda\tilde{\epsilon}(n) + H^T(n)X(n)$

Table I

These RLS equations perform a square root factorization of the autocorrelation matrix. As it is well known, the resulting RLS algorithm is quite robust from a numeric point of view but it requires a lot of computations, being of $O(N^2)$ order.

Although a number of fast RLS algorithms are available in the literature, we adopted the $O(N)$ RLS-SQR algorithm reported in Table II [5]. The main reason of this choice is that the algorithm is more accurate and more robust in limited precision environments with respect to other fast RLS algorithms, as shown experimentally in [5].

1.	$f_{n-1}(n) = v_n + P_{n-1}^T(T_{n-1}^{-1}Z_{n-1})$
2.	$f_n(n) = \gamma_{n-1}f_{n-1}(n)$
3.	$Z_n = T_{n-1}^{-1}Z_{n-1} - W_{n-1}P_{n-1}f_{n-1}(n)$
4.	$\alpha_n = \lambda\alpha_{n-1} + f_n(n)f_{n-1}(n)$
5.	$\sigma_n = \lambda\sigma_{n-1} + v_n^2$
6.	$\begin{pmatrix} P_n \\ b_n(n) \end{pmatrix} = \begin{pmatrix} v_n \\ L_n^T\left(v_n\sigma_n^{-1}W_{n-1}^{-1}Z_n + P_{n-1}\right) \end{pmatrix}$
7.	$\begin{pmatrix} W_n & \mathbf{0} \\ \mathbf{0^T} & \beta_n^{-1} \end{pmatrix} = \begin{pmatrix} \sigma_n^{-1} & \mathbf{0^T} \\ \mathbf{0} & Q_n \end{pmatrix}$
8.	$\gamma_n = \gamma_{n-1} - \alpha_n^{-1}f_n^2(n) + \beta_n^{-1}b_n^2(n)$

Table II

2.2 Stability of the proposed coding algorithm

The coding system includes a quantizer and a feedback loop and hence the overall system is nonlinear and time varying. The issue of BIBO stability of ADPCM systems has recently received great attention [2]. Using ADPCM-LMS models, instability sources have been identified in the overloading of the quantizer [3]. With RLS adaptive prediction, however, the problems associated with stability are much larger. In the following, we consider an ADPCM system where the predictor is updated with an RLS-SQR algorithm (this assumption simplifies the analytical derivations without limiting their applicability).

Proposition 1 *In an ADPCM system with an RLS adaptive prediction algorithm, implemented by means of a square-root approach, an overloading of the quantizer leads to an error of the estimated predictor equal to*

$$\Delta H(n) = \xi_n \frac{L_n}{\beta_n} \qquad (1)$$

at the current time instant n and equal to

$$\Delta H(n+1) = \Delta H(n) + \xi_{n+1}\frac{[L_{n+1} + R_{xx}^{-1}(n)\xi_n]}{[\beta_{n+1} + \xi_n^2 R_{xx}^{-1}(n)_{1,1}]} \qquad (2)$$

at the time instant $n+1$. In the above relations, ξ_n is the perturbation of the estimation error, R_{xx} is the autocorrelation matrix, $L_n = R_{xx}^{-1}(n-1)X(n)$ and $\beta_n = \lambda + X^T(n)R_{xx}^{-1}(n-1)X(n)$, λ being the forgetting factor.

Sketch of the Proof Let us call M the maximum value which can be quantized without overloading. Considering eq7 and eq8 we have, in overloading, $\epsilon_q(n) = M$ and

$$\tilde{\epsilon}(n) = \epsilon_q/\beta(n) = M/\beta(n)$$

The variation of the ϵ coefficient is therefore $\Delta\tilde{\epsilon}(n) = [\epsilon - M]/\beta(n)$ or, letting $\xi_n = \epsilon - M$, we have $\Delta\tilde{\epsilon}(n) = \xi_n/\beta(n)$. From eq9, then, we obtain that the error on the estimated predictor is given by (1). From eq1, it can be concluded that

$$\Delta F(n+1) = S^T(n)[\xi_n 0_{N-1}] = S(n)_1^T\xi_n \qquad (3)$$

Since $R_{XX}^{-1}(n) = S(n)S^T(n)$, from eq2 we have that $\Delta L(n+1) = S(n)\Delta F(n+1) = S(n)S^T(n)[\xi_n 0_{N-1}] = R_{XX}^{-1}(n)_1\xi_n$, where the notation A_i stands for the i-th column of the generic matrix A and With eq3 and (3) we obtain $\Delta\beta(n+1) = \Delta F^T(n+1)\Delta F(n+1) = \xi_n^2 R_{XX}^{-1}(n)_{1,1}a$. Using eq4 and eq5, it can readily shown that

$$\Delta S(n+1) = \alpha(n+1)/\sqrt{\lambda}[\xi_n^2(R_{XX}^{-1}(n)_1 S(n)_1) +$$

$$+\xi_n R_{XX}^{-1}(n)_1 F^T(n+1) + \xi_n L(n+1)S(n)_1]$$

Extending now eq6 as follows: $\epsilon(n+1) = y(n+1) - H^T(n)X(n+1)$, we obtain, using eq8,

$$\Delta\tilde{\epsilon}(n+1) = \xi_{n+1}/[\beta(n+1) + \xi_n^2 R_{XX}^{-1}(n)_{1,1}]$$

Therefore, the final results of eq.(2) is obtained. $\square$

Let us consider now the LMS adaptation rule. It is rather straightforward to show that the following result holds.

Proposition 2 *In an ADPCM system with a normalized LMS adaptive prediction algorithm, an overloading of the quantizer leads to an error of the estimated predictor equal to*

$$\Delta H(n+1) = 2\mu_n \xi_n X(n) \tag{4}$$

at the current time instant n and equal to

$$\Delta H(n+2) = 2\alpha\xi_{n+1}[\frac{(y(n)-\xi_n)X(n)}{(y(n)-\xi_n)^2} + \parallel X(n) \parallel^2] \tag{5}$$

at the time instant $n+1$.

Remark 2 We can note that, for the RLS case, the coefficients' perturbation depends recursively from the perturbation at the previous step. Moreover, the typical values of the parameters in the above equation have been experimentally evaluated, and it was noticed that the perturbation $\Delta H(n)$ is two order of magnitude greater than the same perturbation in the LMS case. In conclusion, the overloading for LMS is much less critical, by at least two order of magnitudes, than the RLS case. This result can also be deduced from the convergence speed difference of the two adaptive algorithms.

The system stabilization problem, hence, is of fundamental importance. In [4], it was introduced a simple stabilization scheme of the RLS algorithm based on a periodic resorting to LMS. We instead used an adaptive switching mechanism based on a minimum error criteria, as shown in Fig.1. The performance obtained with our scheme are very much higher than that obtained using the approach of [4]; in that case, in fact, the RLS section is used only 30% of the time. Using our approach, the RLS algorithm is used at about a 70% rate.

2.3 Sampling Frequency issue

The first experimental observation was that the system performance increases at higher sampling frequencies.

Proposition 3 *If the quantizer of a generic ADPCM system is overloaded, then the following condition holds:*

$$\frac{M_I^S}{f_S} > M \tag{6}$$

where M is the maximum value which can be quantized without overloading, M_I^S is the first order moment of the spectral signal distribution and f_S is the sampling frequency.

Proof The quantizer is overloaded when

$$|e(n)| > (2^{nbit}-1)\Delta/2$$

given that $e(n)$ is the error signal to quantize. Let us suppose that the prediction system was able to accurately follow the signal until the sample index $(n-1)$ and that there is a sudden increase of the error signal. This means that the reconstructed signal didn't track the original signal, but it remained close to the reconstructed signal at the previous time instant. In other words, we can write the following relation:

$$e(n) = s(n) - \hat{s}(n) \approx s(n) - \hat{s}(n-1) \approx s(n) - s(n-1)$$

Moreover, let us assume that the analog signal $s(t)$ be derivable. By turning to the discretization process of $s(n)$, we can say that the time scale is divided into elementary time intervals dt and thet the signal is divided into elementary signals ds. Since $dt = 1/fs$, fs being the sampling frequency, or that $dt \cdot fs = 1$, it comes out that

$$e(n) \approx s(n) - s(n-1) = [s(n) - s(n-1)]/dt \cdot 1/fs =$$

$$= ds/dt \cdot 1/fs$$

Hence the overloading condition is represented by the following condition: $ds/dt \cdot 1/fs > M$. Recalling from the Sampling Theorem that, for a generic signal $x(t)$ with limited bandwidth and finite energy, the following condition holds: $|dx/dt| < M_I^X$, where $M_I^X = \frac{1}{2\pi}\int_{-\omega_n}^{\omega_n} |X(\omega)||\omega|d\omega$ is the first order moment of the spectral signal distribution. We can therefore say that the condition $ds/dt \cdot 1/fs > M$ can be also expressed as $M_I^S/fs > M$. $\square$

Remark 3 From Proposition 2 it can be deduced that the highest is the sampling frequency, the lowest is the probability to overload the quantizer. In other words, the performance increase with the sampling frequency.

2.4 Efficient coding of the side information

The quantization of the side information requires almost 1 bit/sample. However, some schemes for reducing the overhead information sent to the receiver have been devised. The best scheme is summarized as follows:

If the difference between the two reconstructed errors is less than a threshold K_1, or if both the errors are less than another threshold K_2, and if the quantizers are not overloaded, then use the same predictor of the previous time instant. Otherwise, send the a 1 bit information.

Using this scheme, which can be replicated at the receiver, the side info has been reduced down to 0.1 bit/sample, simply by varying the two thresholds; some results are reported in the next section.

3 EXPERIMENTAL RESULTS

The goal of the experiments is to find the best parameter setting, as summarized below:

- optimum prediction orders for LMS and RLS

- optimum coding of the side information (K1, K2 parameters)
- number of quantization bits
- the best sampling frequency

Many experimental verifications and performance measurements of the described algorithm have been performed, using ten ITU's standard test signals which were originally sampled at 48 KHz. The same data set was downsampled at 22.05, 16, 11.025 and 8 KHz, maintaining the highest quality. Some results are reported in Fig.5, where the proposed system is compared with the G.723 standard in order to quantify the increment due to RLS (remember that the LMS part of the proposed algorithm was realized with the G.723 ARMA adaptive predictor algorithm). These results have been obtained with N=15, 3 quantization bits for G.723 and a variable bit-rate side information for the switched system.

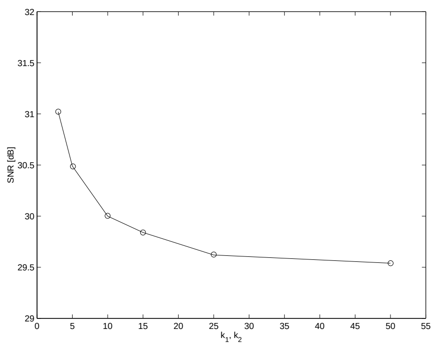

Figure 2: SNR performance vs side coding parameters

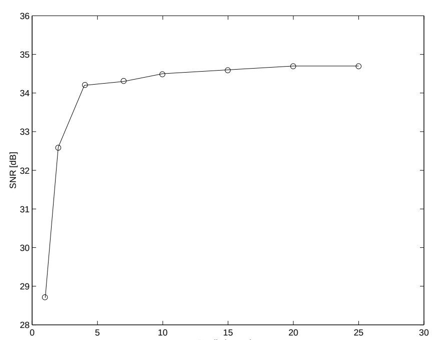

Figure 3: SNR performance at different prediction orders

The performance obtained by varying the coding of the side information (K1, K2 are the thresholds mentioned in II.C) vary smoothly with the sampling rate. This opens to the possibility of using the coding algorithm in embedded coding frameworks.

It is quite important to note that a subjective comparison with the G.722 standard show that the proposed algorithm outperforms the G.722 for all the files of the test-set, with lower computational complexity. The algorithm has been implemented in real-time on

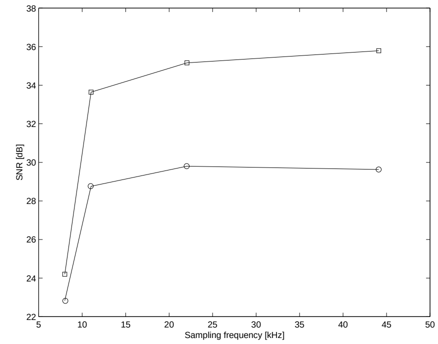

Figure 4: SNR comparison between G723 (circles) and the proposed algorithm (squares) at 5 bit quantization

a 15 MFlops DSP, 320C30 family. The real-time system worked uninterruptly for several days with many different input signals, from tonal and random noise to high quality music, and no stability problem at all arised.

4 Conclusion

In this paper we have described an algorithm for coding wideband signals. Although the performance are quite good, the algorithm should be considered as a starting point for further improvements. The current version, in fact, has no noise masking processing and no particular care in the bit allocation scheme has been given. These two issues will be the following activities on this subject. Thus, we expect a further performance increasing in the near future. The described algorithm has been implemented in real-time in a floating point DSP.

References

[1] M.Dietz, H.Popp, K.Brandemburg, R. Friedrich, “Audio Compression for Network Transmission”, *J. Audio Eng. Soc.*, Vol.44, N.1/2, Jan/Feb 1996

[2] M.Bonnet, O.Macchi, M. Jadaine-Sadaine, “Theoretical Analysis of the ADPCM CCITT Algorithm”, IEEE Trans. on Communic., Vol.38, N.6, Jun 1990

[3] C.R.Watkins, R.R.Bitmead, S.Crisafulli, “Destabilization Effects of Adaptive Quantization in ADPCM”, *IEEE Trans. on Speech and Audio Processing*, Vol.3, N.2, March 95

[4] E.Eleftheriou, D.Falconer, “Tracking Properties and Steady State Performance of RLS Adaptive Filter Algorithms”, *IEEE Trans. on Acoustic, Speech and Signal Processing*, Vol.ASSP-34, N.5, Oct. 1986

[5] A.Carini, E.Mumolo, “Fast Square-Root RLS Adaptive Filtering Algorithms”, *Signal Processing 57*, Elsevier, 233-250, 1997

[6] E.Mumolo, A.Carini,“Volterra Adaptive Prediction of Speech with Application to Waveform Coding”, *Europ. Transaction on Telecomm.*, Vol.6, N.6, Nov. 1995

[7] S. Haykin, “Adaptive Filter Theory”, Prentice-Hall, 1991

[8] N.Bershad, O.Macchi, “Comparison of RLS and LMS Algorithms for Tracking a Chirped Signal”, *Proc. of ICASSP 89*, pp.896-899