Audio Engineering Society
Convention Paper
Presented at the 113th Convention
2002 October 5–8 Los Angeles, CA, USA

Lossless Audio Coding Using Adaptive Multichannel Prediction

Tilman Liebchen

Institute for Telecommunication Systems, Technical University of Berlin, Germany

Correspondence should be addressed to Tilman Liebchen (liebchen@ee.tu-berlin.de)

ABSTRACT
Lossless audio coding enables the compression of digital audio data without any loss in quality due to a perfect reconstruction of the original signal. The compression is achieved by means of decorrelation methods such as linear prediction. However, since audio signals usually consist of at least two channels, which are often highly correlated with each other, it is worthwhile to make use of inter-channel correlations as well. In the current paper it is shown how conventional (mono) prediction can be extended to stereo and multichannel prediction in order to improve compression efficiency. Results for stereo and multichannel recordings are given.

INTRODUCTION

Lossless audio coding techniques [1] [2] [3] [4] [5] often use *Linear Predictive Coding* (*LPC*) to reduce bit rates compared to PCM. Usually the signals are decorrelated by linear prediction, and the residuals are finally entropy coded (Huffman or arithmetic coding).

In conventional LPC systems, individual channels of stereo or multichannel signals are either coded separately (*independent coding*, *mono coding*), or a simple difference coding scheme is applied [3] [4] [5]. Whilst independent coding only removes *intra-channel correlations* from the signals, the bit rates could be further reduced if all channels were jointly coded (*joint stereo coding*), thus also removing *inter-channel correlations*.

Joint stereo methods are common for *lossy* audio coding techniques like MPEG [6] [7] [8]. However, methods such as M/S-Stereo [9] and Intensity Stereo [10] [11] are barely suited for lossless coding, since some irrelevancy is always removed.

The following sections illustrate how the efficiency of lossless audio coding can be increased by using stereo and multichannel prediction. A short outline of linear prediction together with known concepts of exploiting inter-channel correlations is given. This is then extended and generalized to cover multichannel prediction. Since a reasonable choice of predictor *orders* is crucial, appropriate adaptation methods are presented as well. Finally, compression results for different kinds of audio material are given, comparing conventional with stereo and multichannel prediction.

LINEAR PREDICTION

The current sample of a time-discrete signal $x(n)$ can be approximately predicted from its previous values $x(n-k)$. The estimate is given by

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

where K is the order of the predictor. If the predicted values are close to the original samples, the residual

$$e(n) = x(n) - \hat{x}(n) \tag{2}$$

has a smaller variance than $x(n)$ itself, hence $e(n)$ can be encoded more efficiently.

In forward linear prediction, the optimal predictor coefficients a_k (in terms of a minimized variance of the residual) are usually estimated by the autocorrelation method or the covariance method [12]. The autocorrelation method, using the Levinson-Durbin algorithm, has the advantage of providing a simple means to iteratively adapt the *order* of the predictor as well [2].

By increasing the predictor order, the variance of the prediction error will decrease, leading to a smaller bit rate for the residual. On the other hand, the bit rate for the predictor coefficients will rise with the number of coefficients to be transmitted. Thus, the task is to find exactly the order which minimizes the total bit rate.

The Levinson-Durbin algorithm recursively determines all predictors with increasing order. For each order, a complete set of predictor coefficients is calculated. As a side effect, the variance σ_e^2 of the corresponding residual can be calculated as well, resulting in an estimate of the expected bit rate for the residual. Together with the bit rate for the coefficients, the total bit rate can be determined in each iteration, i.e. for each predictor order. When the total bit rate no longer decreases, the order is finally used for prediction.

DIFFERENCE CODING

For stereo signals, it is straightforward to process the two channels $x_1(n)$ (left) and $x_2(n)$ (right) independently. A simple way to exploit dependencies between the channels is to code $x_1(n)$ and the difference

$$d(n) = x_2(n) - x_1(n) \tag{3}$$

instead of $x_1(n)$ and $x_2(n)$. Switching between $x_2(n)$ and $d(n)$ in particular frames depends on which signal can be coded more efficiently. Such prediction with *switchable* difference coding is beneficial in cases where both channels are very similar.

STEREO PREDICTION

Another approach to exploit inter-channel correlations is *stereo prediction* or *stereo LPC* [13]. The principle is shown in Fig. 1. The stereo predictor for each channel uses not only previous samples from the same channel but also samples from the other channel.

Prediction of the left channel

For the left channel $x_1(n)$ the estimate is given by

$$\hat{x}_1(n) = \sum_{k=1}^{K_a} a_k \cdot x_1(n-k) + \sum_{k=1}^{K_b} b_k \cdot x_2(n-k). \tag{4}$$

The first sum represents an *auto-predictor* with order K_a and coefficients a_k, the second sum represents a cross-predictor with order K_b and coefficients b_k. Auto-predictor and cross-predictor must be optimized simultaneously, since they are not independent from each other.

The optimization method is similar to the covariance method in mono prediction.

The variance of the residual $e_1(n) = x_1(n) - \hat{x}_1(n)$ is

$$\sigma_{e_1}^2 = E\left[e_1^2(n)\right] = E\left[(x_1(n) - \hat{x}_1(n))^2\right], \tag{5}$$

which is minimized with respect to the predictor coefficients a_k and b_k:

$$\frac{\partial \sigma_{e_1}^2}{\partial a_k} \equiv 0 \quad , \quad \frac{\partial \sigma_{e_1}^2}{\partial b_k} \equiv 0. \tag{6}$$

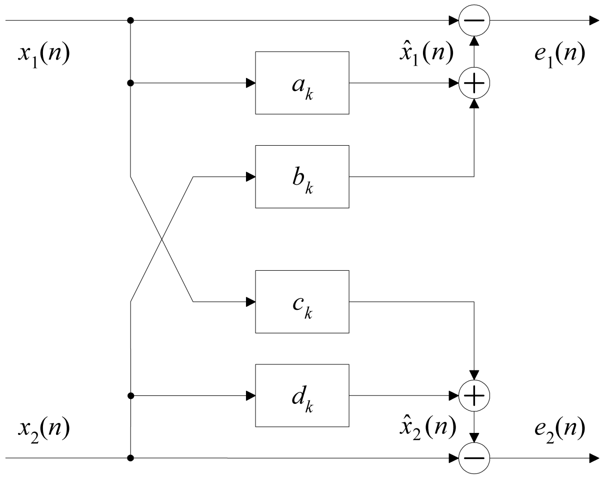

Fig. 1: Principle of stereo prediction

The minimization with respect to a_k gives

$$\frac{\partial \sigma_{e_1}^2}{\partial a_k} = E\left[2 \cdot (x_1(n) - \hat{x}_1(n)) \cdot (-x_1(n-k))\right] = 0 \quad (7)$$

$$E\left[x_1(n) \cdot x_1(n-k)\right] = E\left[\hat{x}_1(n) \cdot x_1(n-k)\right]. \quad (8)$$

By using (4), applying the substitution $l = k$, and by introducing a simplified notation for the autocorrelation

$$r_{pq}(i,j) = r_{x_p x_q}(i,j) = E\left[x_p(n-i) \cdot x_q(n-j)\right], \quad (9)$$

we finally have

$$r_{11}(0,l) = \sum_{k=1}^{K_a} a_k \cdot r_{11}(l,k) + \sum_{k=1}^{K_b} b_k \cdot r_{12}(l,k), \quad (10)$$

where $l = 1 \dots K_a$.

The minimization with respect to b_k gives

$$\frac{\partial \sigma_{e_1}^2}{\partial b_k} = E\left[2 \cdot (x_1(n) - \hat{x}_1(n)) \cdot (-x_2(n-k))\right] = 0 \quad (11)$$

$$E\left[x_1(n) \cdot x_2(n-k)\right] = E\left[\hat{x}_1(n) \cdot x_2(n-k)\right]. \quad (12)$$

Similar to (10) we get

$$r_{12}(0,l) = \sum_{k=1}^{K_a} a_k \cdot r_{21}(l,k) + \sum_{k=1}^{K_b} b_k \cdot r_{22}(l,k), \quad (13)$$

where $l = 1 \dots K_b$.

The merging of (10) and (13) yields the following set of equations

$$\begin{bmatrix} \mathbf{r}_{11} \\ \mathbf{r}_{12} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \mathbf{R}_1 \cdot \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \quad (14)$$

with the correlation vectors

$$\mathbf{r}_{11} = \left[r_{11}(0,1), r_{11}(0,2), \dots, r_{11}(0,K_a)\right]^T, \quad (15)$$

$$\mathbf{r}_{12} = \left[r_{12}(0,1), r_{12}(0,2), \dots, r_{12}(0,K_b)\right]^T, \quad (16)$$

the coefficient vectors

$$\mathbf{a} = \left[a_1, a_2, \dots, a_{K_a}\right]^T, \quad (17)$$

$$\mathbf{b} = \left[b_1, b_2, \dots, b_{K_b}\right]^T, \quad (18)$$

and the correlation matrices

$$\mathbf{R}_{11} = \begin{bmatrix} r_{11}(1,1) & \cdots & r_{11}(1,K_a) \\ \vdots & & \vdots \\ r_{11}(K_a,1) & \cdots & r_{11}(K_a,K_a) \end{bmatrix}, \quad (19)$$

$$\mathbf{R}_{12} = \begin{bmatrix} r_{12}(1,1) & \cdots & r_{12}(1,K_b) \\ \vdots & & \vdots \\ r_{12}(K_a,1) & \cdots & r_{12}(K_a,K_b) \end{bmatrix}, \quad (20)$$

$$\mathbf{R}_{21} = \begin{bmatrix} r_{21}(1,1) & \cdots & r_{21}(1,K_a) \\ \vdots & & \vdots \\ r_{21}(K_b,1) & \cdots & r_{21}(K_b,K_a) \end{bmatrix}, \quad (21)$$

$$\mathbf{R}_{22} = \begin{bmatrix} r_{22}(1,1) & \cdots & r_{22}(1,K_b) \\ \vdots & & \vdots \\ r_{22}(K_b,1) & \cdots & r_{22}(K_b,K_b) \end{bmatrix}. \quad (22)$$

The matrix $\mathbf{R}_1$ has the dimension $(K_a+K_b) \times (K_a+K_b)$. It is symmetric because of

$$r_{21}(i,j) = r_{12}(j,i) \quad (23)$$

and therefore

$$\mathbf{R}_{21} = \mathbf{R}_{12}^T. \quad (24)$$

The coefficients a_k und b_k can be calculated via

$$\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \mathbf{R}_1^{-1} \cdot \begin{bmatrix} \mathbf{r}_{11} \\ \mathbf{r}_{12} \end{bmatrix}. \quad (25)$$

Due to the symmetry of $\mathbf{R}_1$, efficient methods such as Cholesky decomposition can be used.

Prediction of the right channel

We assume that the left channel $x_1(n)$ and the right channel $x_2(n)$ are interleaved, and the samples from each channel are reconstructed alternately in the decoder. In this case, also the *current* sample of $x_1(n)$ can be used to predict $x_2(n)$, and the estimate is given by

$$\hat{x}_2(n) = \sum_{k=0}^{K_c} c_k \cdot x_1(n-k) + \sum_{k=1}^{K_d} d_k \cdot x_2(n-k). \quad (26)$$

The sums represent the cross-predictor and the auto-predictor. The first sum starts at $k = 0$, since the current sample of the left channel is used.

The variance of the residual $e_2(n) = x_2(n) - \hat{x}_2(n)$ is

$$\sigma_{e_2}^2 = E\left[e_2^2(n)\right] = E\left[(x_2(n) - \hat{x}_2(n))^2\right], \tag{27}$$

which is minimized with respect to c_k and d_k:

$$\frac{\partial \sigma_{e_2}^2}{\partial c_k} \equiv 0 \quad , \quad \frac{\partial \sigma_{e_2}^2}{\partial d_k} \equiv 0. \tag{28}$$

Minimization is similar to channel $x_1(n)$. Finally we get

$$r_{21}(0,l) = \sum_{k=0}^{K_c} c_k \cdot r_{11}(l,k) + \sum_{k=1}^{K_d} d_k \cdot r_{12}(l,k), \tag{29}$$

where $l = 0 \dots K_c$, and

$$r_{22}(0,l) = \sum_{k=0}^{K_c} c_k \cdot r_{21}(l,k) + \sum_{k=1}^{K_d} d_k \cdot r_{22}(l,k), \tag{30}$$

where $l = 1 \dots K_d$.

We obtain the following set of equations:

$$\begin{bmatrix} \mathbf{r}_{21} \\ \mathbf{r}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix} = \mathbf{R}_2 \cdot \begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix} \tag{31}$$

with

$$\mathbf{r}_{21} = [r_{21}(0,0), r_{21}(0,1), \dots, r_{21}(0,K_c)]^T, \tag{32}$$

$$\mathbf{r}_{22} = [r_{22}(0,1), r_{22}(0,2), \dots, r_{22}(0,K_d)]^T, \tag{33}$$

and

$$\mathbf{c} = [c_0, c_1, \dots, c_{K_c}]^T, \tag{34}$$

$$\mathbf{d} = [d_1, d_2, \dots, d_{K_d}]^T, \tag{35}$$

as well as

$$\mathbf{R}_{11} = \begin{bmatrix} r_{11}(0,0) & \cdots & r_{11}(0,K_c) \\ \vdots & & \vdots \\ r_{11}(K_c,0) & \cdots & r_{11}(K_c,K_c) \end{bmatrix}, \tag{36}$$

$$\mathbf{R}_{12} = \begin{bmatrix} r_{12}(0,1) & \cdots & r_{12}(0,K_d) \\ \vdots & & \vdots \\ r_{12}(K_c,1) & \cdots & r_{12}(K_c,K_c) \end{bmatrix}, \tag{37}$$

$$\mathbf{R}_{21} = \begin{bmatrix} r_{21}(1,0) & \cdots & r_{21}(1,K_c) \\ \vdots & & \vdots \\ r_{21}(K_d,0) & \cdots & r_{21}(K_d,K_c) \end{bmatrix}, \tag{38}$$

$$\mathbf{R}_{22} = \begin{bmatrix} r_{22}(1,1) & \cdots & r_{22}(1,K_d) \\ \vdots & & \vdots \\ r_{22}(K_d,1) & \cdots & r_{22}(K_d,K_d) \end{bmatrix}. \tag{39}$$

The matrix $\mathbf{R}_2$ has the dimension $(K_c + K_d + 1) \times (K_c + K_d + 1)$. Like $\mathbf{R}_1$, it is symmetric, but now all indices referring to $x_1(n)$ start at zero.

The coefficients can c_k und d_k can be calculated via

$$\begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix} = \mathbf{R}_2^{-1} \cdot \begin{bmatrix} \mathbf{r}_{21} \\ \mathbf{r}_{22} \end{bmatrix}. \tag{40}$$

Thus, (25) und (40) provide all sets of predictor coefficients for both channels.

Adaptation of stereo predictor orders

A crucial aspect of stereo prediction is the choice of suitable predictor orders. The orders of the stereo predictors in [13] are fixed; an adaptation method is not provided.

Contrary to mono prediction, where the predictor coefficients can be calculated using the Levinson-Durbin algorithm, multichannel prediction requires the inversion of the multichannel correlation matrix. Hence, the orders of auto- and cross-predictors have to be known in advance, i.e. prior to applying Cholesky decomposition.

For each channel, two orders have to be specified, but apparently the best orders for auto-predictor and cross-predictor are not independent from each other. Assuming maximum orders $K_a = 30$ (auto-predictor) and $K_b = 10$ (cross-predictor) gives a total of $K_a \cdot K_b = 300$ possible combinations. It is obviously not very efficient to search for the best orders by calculating the resulting bit rates for all of those combinations.

An adaptation method was developed in order to reduce the number of reasonable combinations for each channel. First of all, the optimal order of a mono predictor is determined by using the Levinson-Durbin algorithm. This order is now used as the order for the auto-predictor. A cross predictor is then added. The order of the cross predictor is increased until the bit rate has reached a minimum value. Thus, the correlation matrix has to be inverted only for K_b different cross predictors, reducing the number of matrix inversions by a factor of K_a, or even more if a larger increment of the cross predictor is chosen.

Tests have shown that the optimal order of the auto-predictor is usually somewhat less than the optimal mono order, since high-order intra-channel prediction becomes less important if inter-channel prediction works well. Furthermore, the optimal order of the cross-predictor appeared to be reasonably independent from the auto-predictor's order. Therefore, after adaptation of the cross-predictor, the procedure verifies whether a decrement of the auto-predictor's order leads to an additional bit rate reduction.

MULTICHANNEL PREDICTION

Stereo prediction can be extended to mutual prediction of any number of channels $x_m(n), m = 1 \ldots M$. For a particular channel $x_p(n)$, the estimate is calculated using previous samples from all channels:

$$\hat{x}_p(n) = \sum_{m=1}^{M} \sum_{k=1}^{K_{pm}} a_{pmk} \cdot x_m(n-k). \quad (41)$$

Each partial predictor from channel $x_m(n)$ may have an individual order K_{pm} with predictor coefficients a_{pmk}.

Optimization

The predictor coefficients are calculated using the following approach. First of all, the variance of the residual

$$e_p(n) = x_p(n) - \hat{x}_p(n) \quad (42)$$

is minimized with respect to the predictor coefficients. For all $p = 1 \ldots M$ we set

$$\frac{\partial \sigma_{e_p}^2}{\partial a_{pmk}} \equiv 0. \quad (43)$$

By using (42) and (43) together with (9) we finally have

$$r_{pq}(0,l) = \sum_{m=1}^{M} \sum_{k=1}^{K_{pm}} a_{pmk} \cdot r_{qm}(l,k), \quad (44)$$

where $q = 1 \ldots M$ and $l = 1 \ldots K_{pq}$. Thus, for each channel $x_p(n)$ there are M sets of K_{pq} equations. These equations describe the mutual correlations between all channels and are used to estimate the predictor coefficients.

For each channel $x_p(n)$ we obtain a set of linear equations, which can be written as

$$\begin{bmatrix} \mathbf{r}_{p1} \\ \mathbf{r}_{p2} \\ \vdots \\ \mathbf{r}_{pq} \\ \vdots \\ \mathbf{r}_{pM} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} & \cdots & \mathbf{R}_{1m} & \cdots & \mathbf{R}_{1M} \\ \mathbf{R}_{21} & \mathbf{R}_{22} & \cdots & \mathbf{R}_{2m} & \cdots & \mathbf{R}_{2M} \\ \vdots & \vdots & & \vdots & & \vdots \\ \mathbf{R}_{q1} & \mathbf{R}_{q2} & \cdots & \mathbf{R}_{qm} & \cdots & \mathbf{R}_{qM} \\ \vdots & \vdots & & \vdots & & \vdots \\ \mathbf{R}_{M1} & \mathbf{R}_{M2} & \cdots & \mathbf{R}_{Mm} & \cdots & \mathbf{R}_{MM} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{a}_{p1} \\ \mathbf{a}_{p2} \\ \vdots \\ \mathbf{a}_{pm} \\ \vdots \\ \mathbf{a}_{pM} \end{bmatrix} \quad (45)$$

or

$$\mathbf{r}_p = \mathbf{R}_p \cdot \mathbf{a}_p, \quad (46)$$

with the correlation vectors

$$\mathbf{r}_{pq} = [r_{pq}(0,1), r_{pq}(0,2), \ldots, r_{pq}(0,K_{pq})]^T, \quad (47)$$

the coefficient vectors

$$\mathbf{a}_{pm} = \left[a_{pm1}, a_{pm2}, \ldots, a_{pmK_{pm}}\right]^T, \quad (48)$$

and the correlation matrices

$$\mathbf{R}_{qm} = \begin{bmatrix} r_{qm}(1,1) & \cdots & r_{qm}(1,K_{pm}) \\ \vdots & & \vdots \\ r_{qm}(K_{pq},1) & \cdots & r_{qm}(K_{pq},K_{pm}) \end{bmatrix}. \quad (49)$$

Each partial matrix $\mathbf{R}_{qm}$ has a dimension $K_{pq} \times K_{pm}$, depending on the individual predictor orders. The multichannel correlation matrix $\mathbf{R}_p$ for the prediction of channel $x_p(n)$ from all channels has the dimension $K_p \times K_p$ with

$$K_p = \sum_{m=1}^{M} K_{pm}, \quad (50)$$

which is also the dimension of $\mathbf{r}_p$ and $\mathbf{a}_p$.

Since $\mathbf{R}_p$ is symmetric, (46) can be solved efficiently by the use of Cholesky decomposition, which is equivalent to

$$\mathbf{a}_p = \mathbf{R}_p^{-1} \cdot \mathbf{r}_p, \quad (51)$$

where $\mathbf{a}_p$ contains all sets of predictor coefficients $\mathbf{a}_{pm}$ for channel $x_p(n)$.

Prediction based on current samples

So far only *previous* samples of all channels have been used for the prediction of each channel. As shown for stereo prediction, *current* samples of other channels can also be used, if they are available for the decoder prior to the sample being reconstructed. We assume interleaving of channels, which means that, for all n, sample $x_p(n)$ is reconstructed before $x_{p+1}(n)$ but after $x_{p-1}(n)$. Hence (41) has to be slightly modified, and we have

$$\hat{x}_p(n) = \sum_{m=1}^{M} \sum_{\substack{k=1, p \le m \\ k=0, p>m}}^{K_{pm}} a_{pmk} \cdot x_m(n-k). \quad (52)$$

In this case, (46) is still valid, but the dimensions of the (partial) matrices and vectors have increased, since elements such as $r_{pq}(0,0)$ have to be taken into account.

Thus, the correlation vectors $\mathbf{r}_{pq}$ with $q < p$ are

$$\mathbf{r}_{pq} = [r_{pq}(0,0), r_{pq}(0,1), r_{pq}(0,2), ..., r_{pq}(0,K_{pq})]^T, \quad (53)$$

the coefficient vectors $\mathbf{a}_{pm}$ with $m < p$ are

$$\mathbf{a}_{pm} = \left[a_{pm0}, a_{pm1}, a_{pm2}, ..., a_{pmK_{pm}}\right]^T, \quad (54)$$

and the correlation matrices $\mathbf{R}_{qm}$ with $q < p$ and $m < p$ are

$$\mathbf{R}_{qm} = \begin{bmatrix} r_{qm}(0,0) & \cdots & r_{qm}(0,K_{pm}) \\ \vdots & & \vdots \\ r_{qm}(K_{pq},0) & \cdots & r_{qm}(K_{pq},K_{pm}) \end{bmatrix}. \quad (55)$$

This means that the dimensions of these matrices are increased by one due to the new sum indices.

We should also observe those matrices $\mathbf{R}_{qm}$ where $q \geq p$ or $m \geq p$. For $q \geq p, m < p$ we have

$$\mathbf{R}_{qm} = \begin{bmatrix} r_{qm}(1,0) & \cdots & r_{qm}(1,K_{pm}) \\ \vdots & & \vdots \\ r_{qm}(K_{pq},0) & \cdots & r_{qm}(K_{pq},K_{pm}) \end{bmatrix}, \quad (56)$$

and for $q < p, m \geq p$ we have

$$\mathbf{R}_{qm} = \begin{bmatrix} r_{qm}(0,1) & \cdots & r_{qm}(0,K_{pm}) \\ \vdots & & \vdots \\ r_{qm}(K_{pq},1) & \cdots & r_{qm}(K_{pq},K_{pm}) \end{bmatrix}. \quad (57)$$

In (56) the number of columns is increased, whereas in (57) the number of rows is increased.

For $q \geq p$ and $m \geq p$, the correlation matrix is still defined by (49), as it is without the use of current samples.

Adaptation of multichannel predictor orders
The adaptation method described for stereo prediction can be extended to more than two channels, but in such cases the dependencies between the individual orders might become much more complicated, depending on the audio material.

While the orders of all auto-predictors might be determined using the Levinson-Durbin algorithm, the optimum orders of the cross-predictors strongly depend on the configuration of the channels and the relations between them. For example, 5.1-channel surround material will normally contain stronger correlations between the two front channels than between the center and the rear channels. In such cases, some of the cross-predictors might become obsolete, reducing the total number of reasonable combinations.

RESULTS

Several audio CDs (16 bit, 44.1 kHz, stereo) have been coded using stereo prediction together with the described adaptation method. The maximum orders of the auto-predictor and the cross-predictor were set at 30 and 10 respectively. According to (26), the right channel was predicted using the current sample from the left channel.

Results for independent coding and switchable difference coding, as described in the corresponding section, were calculated for means of comparison. For both methods, the order of the predictor was adapted using the Levinson-Durbin algorithm, but also with a maximum value of 30.

In all cases, an adaptation block length of 1024 was chosen, and 12 bits per predictor coefficient were used. The residuals were entropy coded using Rice codes [2] [3].

Table 1 shows the results for independent coding (*Mono*), difference coding (*Diff*) and stereo coding (*Stereo*).

CD	Mono	Diff	Stereo
Berlioz [14]	6.61	6.61	6.57
Carreras [15]	8.12	7.67	7.62
Chapman [16]	9.08	8.58	8.54
Springsteen [17]	9.98	9.93	9.77
U2 [18]	10.05	9.81	9.67
Vega [19]	9.13	8.85	8.64
Vivaldi [20]	7.08	7.06	6.96

Table 1: Coding results (bits per sample) for different audio CDs (averaged over individual tracks, without silence)

Difference coding achieves savings of up to 0.5 bps, compared with independent coding (Chapman, Carreras). For this material, stereo coding usually yields a small additional saving of approximately 0.05 bps on average and up to 0.1 bps for individual tracks. Similar additional savings are achieved even on material which can not be further compressed by difference coding (Berlioz).

For all other material (Springsteen, U2, Vega, Vivaldi), there are savings of a further 0.1 – 0.2 bps on top of the savings made by difference coding. For some tracks even savings of 0.3 bps and above have been measured.

Recording	Mono	Multichannel
Brubeck [21]	5.80	5.78
Lang [22]	6.25	6.08
Mahler [23]	5.30	5.30
Young [24]	5.35	5.30

Table 2: Coding results (bits per sample) for different multichannel recordings (16 bit, 48 kHz, averaged over extracts from all tracks)

Moreover, some multichannel tracks with six channels (L, R, Ls, Rs, C, LFE), which were recorded from the analog outputs of a DVD-Audio/SACD player with 16 bit, 48 kHz, have been coded using multichannel prediction according to (52). The maximum order of each auto-predictor was set at 20, whilst the maximum order of all cross-predictors was set at 10. In order to keep the adaptation of the cross-predictors simple, *all* orders were increased simultaneously.

Table 2 shows the results for independent coding and multichannel coding. For some material (Lang), the savings are around 0.2 bps on average, and up to 0.3 bps for particular tracks. The individual results for each channel indicated that significant correlations exist between the two front channels, as well as between the two rear channels. However, compared to stereo material, the savings for many tracks are rather small. This might be due to the fact that the recordings have been mixed very spacially.

CONCLUSION

Lossless audio coding can be improved using adaptive stereo prediction. Compared to conventional lossless audio coding techniques, considerable savings are achieved for most stereo signals.

Multichannel prediction yields acceptable savings for some recordings. The results might be improved by the use of a more elaborate algorithm for the adaptation of the predictor orders. Without doubt, multichannel prediction requires further investigation.

REFERENCES

[1] C. Cellier, P. Chenes, and M. Rossi, "Lossless Audio Data Compression for Real Time Apllications," *95th AES Convention*, 1993.

[2] T. Robinson, "SHORTEN: Simple lossless and near-lossless wavform compression," *Technical report CUED/F-INFENG/TR.156, Cambridge University Engineering Department*, 1994.

[3] A. A. M. L. Bruekers, A. W. J. Oomen, R. J. van der Vleuten, and L. M. van de Kerkhof, "Lossless Coding for DVD Audio," *101st AES Convention*, 1996.

[4] P. Craven and M. Gerzon, "Lossless Coding for Audio Discs," *J. Audio Eng. Soc.*, vol. 44, no. 9, September 1996.

[5] M. Hans and R. W. Schafer, "Lossless Compression of Digital Audio," *IEEE Signal Processing Magazine*, July 2001.

[6] J. Johnston, J. Herre, M. Davis, and U. Gbur, "MPEG-2 NBC Audio - Stereo and Multichannel Coding Methods," *101st AES Convention*, 1996.

[7] M. Bosi et al., "ISO/IEC MPEG-2 Advanced Audio Coding," *J. Audio Eng. Soc.*, vol. 45, no. 10, October 1997.

[8] P. Noll, "MPEG Digital Audio Coding - Setting the Standard for High-Quality Audio Compression," *IEEE Signal Processing Magazine*, September 1997.

[9] A. Ferreira J. Johnston, "Sum-Difference Stereo Transform Coding," *Proc. IEEE ICASSP*, 1992.

[10] R. van der Waal and R. Veldhuis, "Subband Coding of Stereophonic Digital Audio Signals," *Proc. IEEE ICASSP*, 1991.

[11] J. Herre, K. Brandenburg, and D. Lederer, "Intensity Stereo Coding," *96th AES Convention*, 1994.

[12] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*, Prentice-Hall, 1984.

[13] P. Cambridge and M. Todd, "Audio Data Compression Techniques," *94th AES Convention*, 1993.

[14] Hector Berlioz, *Symphonie Fantastique*, Chicago Symphony Orchestra, Claudio Abbado, Polydor, 1984.

[15] Carreras, Domingo, and Pavarotti, *In Concert*, Zubin Metha, Decca, 1990.

[16] Tracy Chapman, *Tracy Chapman*, Elektra, 1988.

[17] Bruce Springsteen, *Born in the U.S.A.*, CBS, 1984.

[18] U2, *Achtung Baby*, Island, 1991.

[19] Suzanne Vega, *Solitude Standing*, A&M, 1987.

[20] Antonio Vivaldi, *The Four Seasons*, Nigel Kennedy, English Chamber Orchestra, EMI, 1989.

[21] The Dave Brubeck Quartet, *Time Out*, SACD, Columbia, 1997.

[22] K. D. Lang, *Invincible Summer*, DVD-A, Warner, 2000.

[23] Gustav Mahler, *Symphony No. 10*, DVD-A, Berliner Philharmoniker, Simon Rattle, EMI, 2001.

[24] Neil Young, *Road Rock*, DVD-A, Warner, 2000.