EXHIBIT U

TO DECLARATION OF S. MERRILL WEISS IN SUPPORT OF PLAINTIFF ACACIA MEDIA TECHNOLOGIES CORPORATION'S MEMORANDUM OF POINTS AND AUTHORITIES IN OPPOSITION TO ROUND 3 DEFENDANTS' MOTION FOR SUMMARY JUDGMENT OF INVALIDITY UNDER 35 U.S.C. § 112 OF THE '992, '863, AND '702 PATENTS; AND SATELLITE DEFENDANTS' MOTION FOR SUMMARY JUDGMENT OF INVALIDITY OF THE '992, '863, AND '720 PATENTS

Preprint 2813 (E-6)

Study and evaluation of new method of ADPCM encoding

Yukio Takahashi, Hiroyuki Yazawa,
Kaoru Yamamoto, Takeaki Anazawa

Nippon Columbia Co., Ltd.
Tokyo, Japan

Presented at the 86th Convention 1989 March 7-10 Hamburg

AES

AN AUDIO ENGINEERING SOCIETY PREPRINT

Study and Evaluation of New Method of ADPCM Encoding

Yukio Takahashi, Hiroyuki Yazawa, Kaoru Yamamoto, Takeaki Anazawa
Nippon Columbia Co.,Ltd. Recording and Engineering Department

ABSTRACT

Vulnerable points of ADPCM encoding after CD-I format have been studied. A new processing structure of encoding is proposed as a solution for the characteristic noise and the deterioration of sound quality, and its performance is examined through the result of evaluation test. Consideration about general problems in the realization of ADPCM encoder is also mentioned.

1. INTRODUCTION

Compact Discs use PCM with a sampling frequency of 44.1 kHz and 16-bit linear quantization. In addition to this standard, CD-I also provides for three other formats with differing bandwidth and compression ratio, as shown in Table 1. The BRR (Bit rate reduction) encoding / decoding system has many advantages, such as good sound quality, playing time and simple structure of decoder, etc.. But during Level B and C processing, the characterictic noise or the deterioration of sound quality sometimes occurs according to the spectral and level distribution of input signal, and processing structure of encoder.
Several of these problems can be solved by improving processing structure.

2. BRR SYSTEM

Fig.1 shows the blockdiagram of BRR system. In the following equation, the encoder input is $x(n)$, the prediction error $d(n)$, the quantization error $e(n)$, the encoder output $\hat{d}(n)$, the decoder input $\hat{d}'(n)$, and the decoder output $\hat{x}'(n)$. After Z transformation, each of these value becomes $X(z)$, $D(z)$, $E(z)$, $\hat{D}(z)$, $\hat{D}'(z)$, and $\hat{X}'(z)$, respectively. Assuming that there is no error between encoder output and decoder input, the following equation can be applied

$$\hat{X}'(z) = X(z) + G^{-1}\cdot E(z)\ \{1-R(z)\}/\{1-P(z)\} \quad (1)$$

Where $P(z)$,$R(z)$ are the transfer functions of prediction circuit and noise shaper. Ordinarily, they are set as $P(z) = R(z)$.
G is the gain (shift amount) used for normalizing the prediction error with the maximum value in the block (28 samples). It is defined as

$G = 2^{8-R}$ (Level A), $G = 2^{12-R}$ (Level B and C). R is range data.

The transfer function of prediction filter is $H(z) = 1-P(z)$. The CD-I Audio formats provide for four types (Fig.2), which are selected for each block depending on the spectral content of the input signal. Filter data (Filter type no.)are transfered once in a block combined with range data. The simplest algolithm to achieve prediction adaptation is shown in Fig.3. At the output of each of the four filters, a memory circuit with a storage capacity of one block and a peak hold circuit are used. The filter with the lowest peak within a block is then selected as a filter for the block, and the memory data are output at the next block.

3. ENCODING PROCESS

This section describes actual BRR processing with regard to problems and possible solutions.

(1) Influence of saturation in shifter

When digital data processing of Level B or C is carried out with 16-bit-word length, the decoder output sometimes contains unacceptable noise in case of the encoder input signals with sharp transients, such as trumpet solo. Especially when shifter gain G changes rapidly, the noise level tends to increase.
After study of this phenomenon, we found out it was due to saturation in the shifter of encoder's block-floating part,as described below.

The level distribution range of the prediction filter output d(n) in Fig.1 is as follows (the calcuration is carried out with a 2's compliment).

$$G^{-1}\cdot Nmax \leq d(n) \leq G^{-1}\cdot Pmax \qquad Nmax = 8000(hex),\ Pmax = 7FFF(hex)$$

The adder circuit output d'(n) is as below

$$d'(n) = d(n) - \tilde{e}(n)$$

Therefore, the level distribution range of d'(n) is wider than that of d(n). As the shifter gain G of the next stage is determined by the maximum value of d(n) in the block, the shifter output d"(n) for 16-bit processing does not exceed regardless of the value of d'(n) as following equation

$$Nmax \leq d''(n) \leq Pmax$$

It follows that when d'(n) has a higher value than the possible maximum for d(n), d"(n) saturates at Nmax or Pmax.
In this condition, we can consider that multiplication factor a(n) is used for e(n) as Fig.4, where

$$G^{-1}\cdot Nmax \leq d(n) - a(n)\tilde{e}(n) \leq G^{-1}\cdot Pmax$$

So the transfer function between encoder input and decoder output becomes

$$\hat{X}'(z) = X(z) + G^{-1}\cdot E(z)\ \{1-A(z)*R(z)\}/\{1-P(z)\}$$

Therefore the quantization error has uneven spectrum, which increase the noise level in the decoder output. This effect is especially marked when the shifter gain G increases (input signal level falls). This is because the relative level of the noise shaper output e(n) vs ,the prediction error d(n) increases, thereby increasing the chance of shifter saturation.

Various countermeasures for this phenomenon are possible. The easiest one is changing noise shaper algorithm as shown in Fig.5. When we express the noise shaper input and its Z transformation as e'(n) and E'(z), then

$$G^{-1}\cdot\hat{d}(n) = d(n)-\tilde{e}(n)+e'(n) \qquad (5)$$

$$\hat{d}(n) = G\{\ d(n)-\tilde{e}(n)+e'(n)\ \}$$

The encoder output and the decoder output are

$$\hat{D}(z) = G \cdot X(z)\{1-P(z)\} + G \cdot E'(z)\{1-R(z)\} \qquad (6)$$

$$\hat{X}'(z) = X(z) + E'(z)\{1-R(z)\}/\{1-P(z)\} \qquad (7)$$

The relationship between the quantization error E(z) and E'(z) can be expressed as follows

$$E'(z) = G^{-1} \cdot E(z) \qquad (8)$$

In this condition, provided that d(n) is not close to the maximum 16-bit value, the quantization error is correctly fed back via the noise shaper even if d"(n) saturates. Therefore no irregular distribution occurs in the noise spectrum. Tests have shown that the unacceptable noise described above does not occur, also when the shifter gain fluctuates considerably. Fig.6 shows the decoder output when a square wave signal is fed to the encoder on Level B processing. (a) is the waveform of decoder output with the unchanged algorithm, and (b) is the waveform after the alteration.

(2) Influence of unsymmetrical quantization

Due to the 2'compliment, the code has a unsymmetrical positive/negative distribution. In terms of absolute value, the maximum positive value is smaller than the maximum nagative value by 1LSB. With 16-bit quantization this has almost no influence, but when the number of bits is lower, such as when using the bit rate reduction, the importance of 1 LSB increases and effect cannot be disregarded.
The distribution range of the quantization error e(n) is different, depending on whether the input signal is negative or positive. When R(Z) equal zero, d'(n) can equals d(n), and the quantization step Δq gives

When $d'(n) < 0, \quad -1/2 \cdot \Delta q < e(n) < +1/2 \cdot \Delta q$

When $d'(n) \geq 0, \quad -1/2 \cdot \Delta q < e(n) < +1/2 \cdot \Delta q + b$

The parameter b is determined by the level of d(n) and the number of quantization bits. For example in case of 4 bit, the following applies.

$b = 0$ when $0 \leq |d'(n)| \leq 7$ (hex)

$0 < b < 1/2 \cdot \Delta q$ when $3FFF \leq |d'(n)| \leq 7FFF$

Inserting equation (5) gives

$$c(n) = G^{-1} \cdot \hat{d}(n) - d(n) = e'(n) - \tilde{e}(n)$$

and Z transformation of both members gives

$$C(z) = E'(z)\{1-R(z)\}$$

Therefore $E'(z) = C(z)/\{1-R(z)\}$ (6)

If the peak detector circuit in Fig.3 simply determines G based on the absolute peak value of the prediction filter output, the phenomenon described in the following paragraph occurs, depending on the condition of the input signal.
When x(n) is a low frequency signal and d(n) continuously is close to the positive saturation level of the shifter, c(n) is

$$-G^{-1} \Delta q < c(n) < 0$$

So it is continuously negative, and c(n) contains many mid-to-low frequency components. When prediction filter type 2 or 3 is selected,

the mid-to-low frequency amplification is as shown by equation (6). Therefore $e'(n)$ increases and the decoder output waveform becomes distorted. Fig.7 shows an example for this kind of distortion in the decoder output with a 50Hz sine wave signal input on Level B processing. In this case ,better result can be obtained when the G is reduced by 1. To reduce the influence of this effect upon sound quality, the following strategy is effective.
On a range detection circuit in Fig.8, the output of prediction filter is devided into negative and positive components. The positive output is multiplied with a constant higher than 1 before being supplied to the peak hold circuit. This raises the saturation level of the shifter for positive signal components, thereby providing compensation for the unsymmetrical quantization characteristics.

3.ADPCM real-time encoder

We developped ADPCM real-time encoder(DENON DN-060) which involves new algorithms described above. The unit has following features.
(1) Performs ADPCM encoding (lebel A, B and C) in real-time for up to four channels and supplies data according to the CD-I audio sector format,to an external PCM processor via an AES/EBU interface.
(2) Incorporates a real-time decoder for monitoring during program production.
(3) RS-232C interface permits setting of encoding mode, etc..

Fig.9 shows block diagram of the unit. Music signals are processed as follows.

(1) Encoding process

The signal is input to the A/D converter of the encoder, where it is digitized with a sampling frequency of 37.8kHz or 18.9kHz and 16 bit linear quantization. In the next stage, it is compressed to the Level specified (A,B,or C) by the DSP. These data are temporarily stored in the sector memory, the bits are rearranged to fit the CD-I Audio sector format and the sector interleave is applied. The formatter then attaches sync, header and subheader,and the resulting data are output via AES/EBU interface.

(2) Decoding process

When CD-I audio format signal is supplied to the decoder via AES/EBU interface, the sector memory and DSP perform opposite processing to restore the 16-bits information. The D/A converter then transforms the signal back into analog audio signal.

4.Evaluation test

Using the CD-I audio encoder, a sound quality evaluation was carried out.

(1) Sources where considerable sound quality degradation occurs.

After improvement of encoding algorithm, ADPCM processing still has some problems for some source signal types. Due to the principles of ADPCM processing, the degree of sound quality degradation depends considerably on spectral distribution of the music signal. Especially with Level B and C processing, S/N ratio deteriorates noticeably when sources with a high amount of high frequency energy are used(like brass,bells and piano, etc.) Degradation is most noticeable when these instruments are played as solo, but when they appear in combination with other sources containing many low frequency component, it is much less apparant.

(2) Comparison to other media

To evaluate the ADPCM sound quality, a comparison was also made to other audio media for Level A and B. As preliminary test, popular music and classic music sources were used for general comparison. With the popular music sources, hardly any difference could be found between CD, Level A and Level B. With classic music source on the other hand, the differences were quite easily discernible by ear.

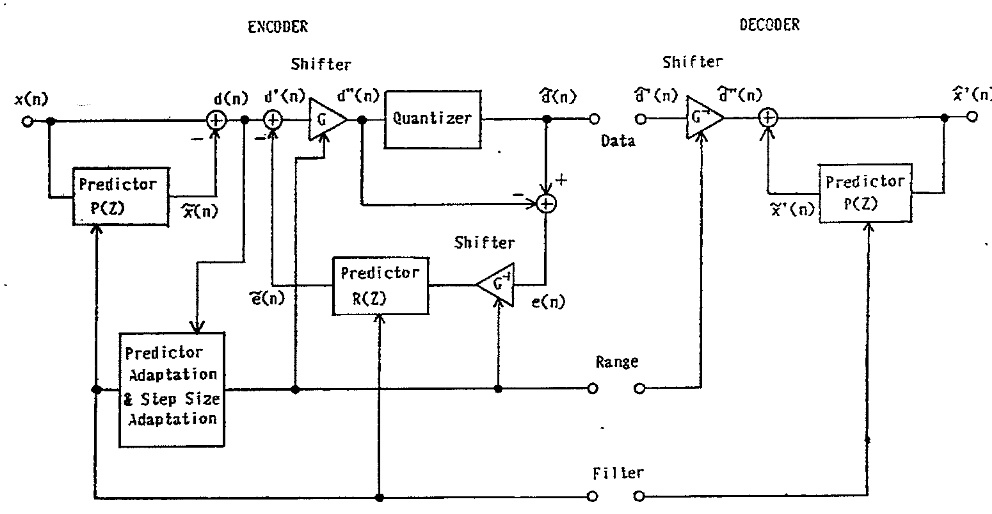
In the main test, CD, cassette, ADPCM Level A and Level B were compared, using piano and orchestra music as sources in a series of one-to-one comparisons to determine a ranking. The result is shown Fig.10. For piano (a), the order was CD, Level A, Level B, cassette. Whereas the difference between CD and Level A was minor, the difference between Level A and Level B as well as Level B and cassette were quite distinct. The low ranking of Level B and cassette can probably be attributed to quantization noise with the former and tape hiss with the latter. With orchestra music (b), there was no difference between CD and Level A, and the difference among the other formats were also less pronounced. This was probably partly due to the fact that a forte passage was used for listening evaluation, making difference in S/N ratio less apparent.

5.Conclusion

Improvements to the processing algorithm for CD-I ADPCM encoding were effective, in reducing the deterioration of sound quality. The evaluation of sound quality showed that level A can be used for "foreground"(serious) music applications, whereas level B is well suited for BGM listening. Depending on the application, level B may also used as "foreground" music medium.
Further reseach into filter selection algorithm and other aspects of ADPCM processing is bound to reduce sound quality degradation more.

References

1.M. Nishiguchi, K. Akagiri and T. suzuki, "A New Audio Bit Rate Reduction System for the CD-I format," presented at the 81st Convention of Audio Engineering Society (November. 1986), preprint 2375.



$x(n)$; encoder input $d(n)$; prediction error
$\tilde{x}(n)$; predicted value $\hat{d}(n)$; encoder output
$e(n)$; quantization error
$\tilde{e}(n)$; predicted value of quantization error

$\hat{d}'(n)$; decoder input
$\tilde{x}'(n)$; predicted value
$\hat{x}'(n)$; decoder output

Figure 1 Block diagram of the bit rate reduction system

Level	sampling Frequency	number of bits / sample	bandwidth	maximum number of Channel
PCM	44.1 kHz	16	20 kHz	1 stereo
ADPCM Level A	37.8 kHz	8	17 kHz	2 stereo 4 mono
Level B	37.8 kHz	4	17 kHz	4 stereo 8 mono
Level C	18.9 kHz	4	8.5 kHz	8 stereo 16 mono

Table 1 Overview of Sound Quality Levels

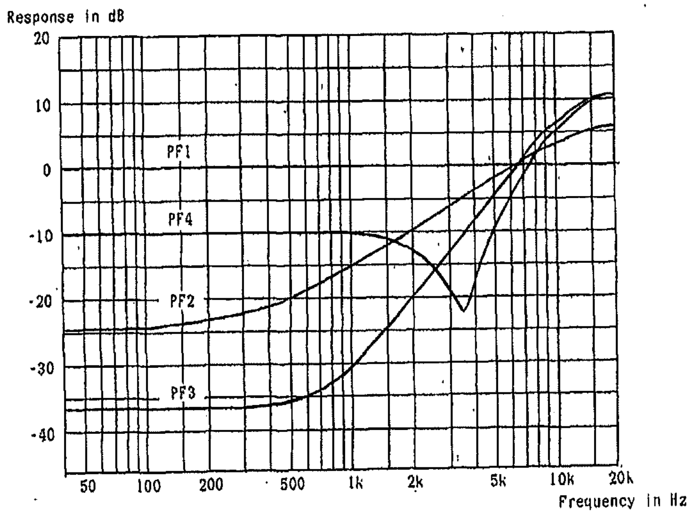


Figure 2 Frequency response of prediction filter when fs = 37.8 kHz
PF1 ; $H(z) = 1$
PF2 ; $H(z) = 1 - 0.9375 z^{-1}$
PF3 ; $H(z) = 1 - 1.796875 z^{-1} + 0.8125 z^{-2}$
PF4 ; $H(z) = 1 - 1.53125 z^{-1} + 0.859375 z^{-2}$

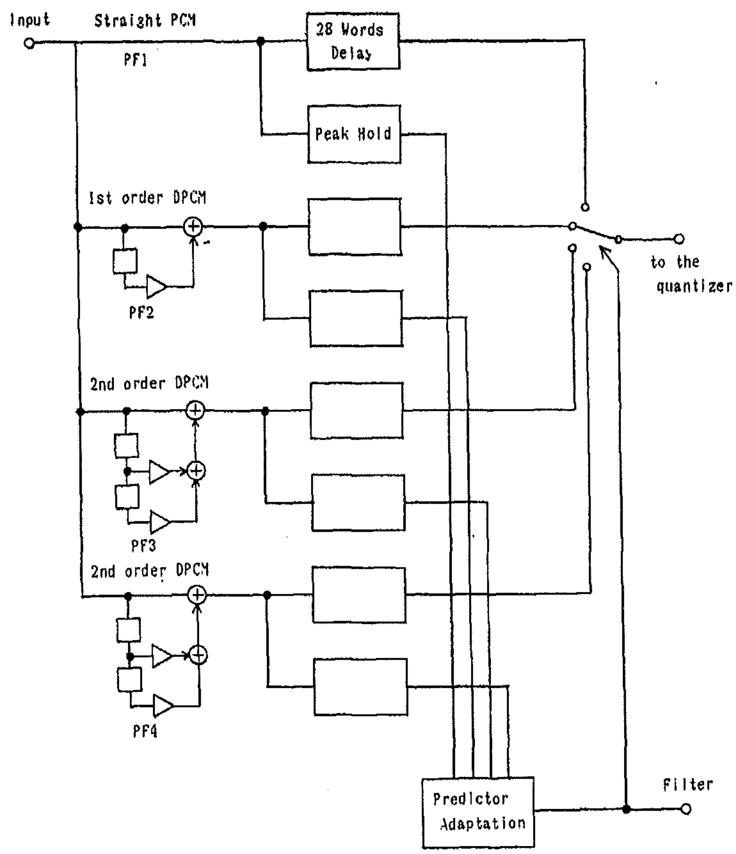


Figure 3 The simplest algolithm of predictor adaptation

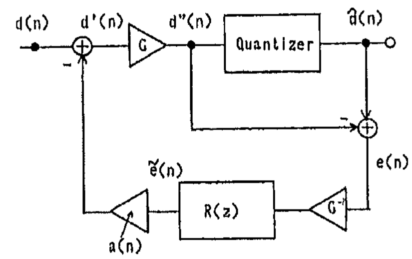


Figure 4 Equivalent circuit of block floating stage when shifter saturation occurs

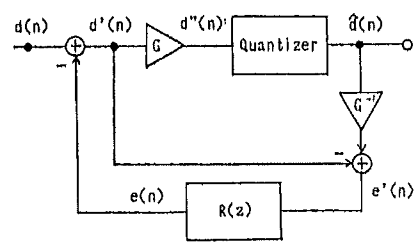


Figure 5 Altered circuit of block floating stage

(a)

(b)

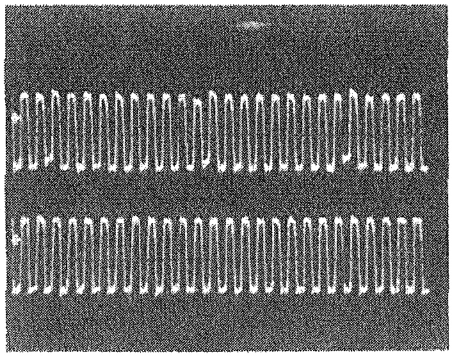


Figure 6 Example of unacceptable noise on decoder output
due to shifter saturation for 1.5kHz square wave
(a) unchanged algolithm with 16 bits system
(b) altered algolithm with the same system
on Level B processing.

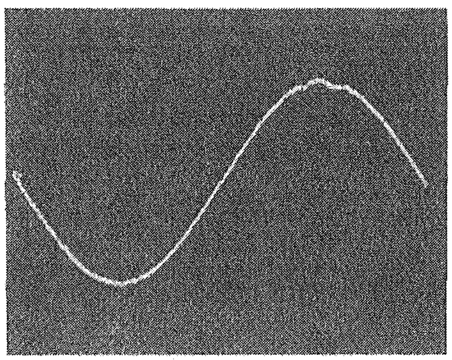


Figure 7 Example of influence of unsymmetrical quantization
(50Hz sine wave)
on Level B processing.

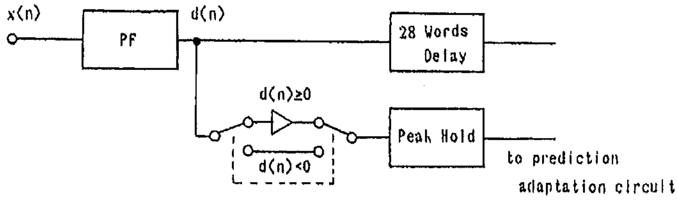


Figure 8 Altered range detector

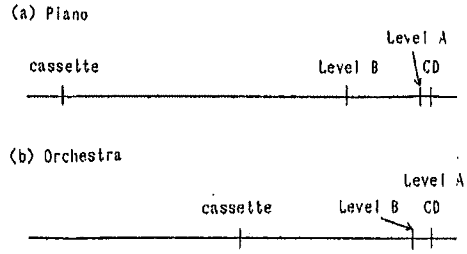


Figure 10 The result of evaluation test

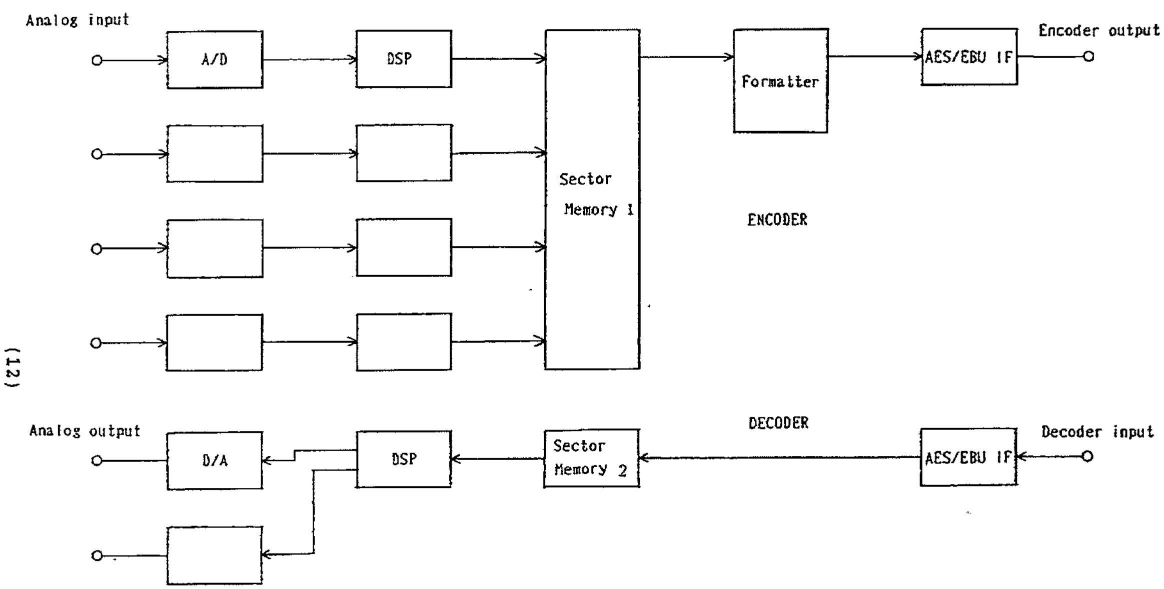


Figure 9 Block diagram of DENON CD-1 Realtime AUDIO ENCODER