# Extended Pilot-Based Coding for Lossless Bit Rate Reduction of MPEG Surround

Hee-Suk Pang, Jaehyun Lim, and Hyen-O Oh

*ABSTRACT—Pilot-based coding (PBC), which is used for lossless bit rate reduction of audio coding, has been recently proposed for MPEG Surround. We propose extended PBC for further lossless bit rate reduction of MPEG Surround. Extended PBC selects the number of pilots depending on the parameter band number and the type of spatial parameter. It then encodes the pilots and the relevant difference data. Experiments show that extended PBC is more effective than the original PBC, especially for high bit rate modes, with a negligible complexity increase on the decoder side.*

*Keywords—Pilot-based coding, lossless bit rate reduction, MPEG Surround.*

## I. Introduction

The simplest way to compress multi-channel signals is to compress each channel sound, which guarantees a good sound quality at the cost of bit rates–for example, 64 kbps for each channel for Advanced Audio Coding (AAC) [1]. Another method of multi-channel audio coding is based on matrix-based downmix and upmix processes [2]. Though the bit rates of the processes are as low as those of normal stereo audio coding, the processes produce poor sound quality, which is caused by information losses in the downmix.

By extracting additional information from original multi-channel signals and embedding it in the bit stream, binaural cue coding (BCC) achieves both a low bit rate and good sound quality [3], [4]. Based on the idea of BCC, MPEG Surround, which is currently in the standardization process, uses spatial parameters composed of channel level difference (CLD), inter-channel correlation/coherence (ICC), and a channel prediction coefficient (CPC) to reconstruct multi-channel sounds from the downmixed sound [5]. Since the bit rate of MPEG Surround in normal mode is not more than 160 kbps including an AAC-encoded stereo signal and spatial parameters, it is adequate for streaming services such as multi-channel digital multimedia broadcasting (MC-DMB).

Recently pilot-based coding (PBC) has been proposed and adopted for lossless bit rate reduction of MPEG Surround [6], [7]. This letter proposes extended PBC for further bit rate reduction of MPEG Surround.

Manuscript received Sept. 11, 2006; revised Oct. 09, 2006.

Hee-Suk Pang (phone: + 82 2 526 4371, email: hspang@lge.com) and Jaehyun Lim (email: jhlim@lge.com) are with DM Research Lab., LG Electronics, Seoul, Korea.

Hyen-O Oh (email: genioho@lge.com) is with DTV Research Lab., LG Electronics, Seoul, Korea.

## II. Brief Review of Pilot-Based Coding

This section presents a brief review of pilot-based coding in MPEG Surround [6], [7]. Suppose that a data set $x(n)$ is composed of $N$ samples. PBC first determines a pilot as

$$p = \arg\min_{p} \sum_{n=1}^{N} g(x(n) - p), \qquad (1)$$

where $p$ is the pilot, $g(x)$ is the number of bits for representation of $x$, and $p$ is generally around the mean value of the set. The PBC encodes $N$+1 values, the pilot and difference data from the pilot, for bit stream composition.

MPEG Surround bit streams are composed of the index values of spatial parameters. In an MPEG Surround encoder, PBC, grouped PCM coding, and time/frequency differential coding are applied to find the best coding method as shown in Fig. 1(a). This procedure is on a parameter set basis, except for time pair coding which is based on two parameter sets. In an MPEG Surround decoder, the bit stream is parsed and checked to determine the coding method as shown in Fig. 1(b). For PBC, a pilot and the difference data are read and then the index values are subsequently calculated as

$$x(n) = p + x_d(n), \quad (2)$$

where $p$ and $x_d(n)$ are the pilot and the difference data. A one-bit flag is used to discriminate PBC from grouped PCM coding.

When the parameter band number is too small, PBC is not advantageous due to additional bit consumption for the pilot. For MPEG Surround, PBC is applied only when the parameter band number is equal to or greater than 5. Since PBC shares Huffman codebooks with the differential coding, it does not require any additional codebook tables.

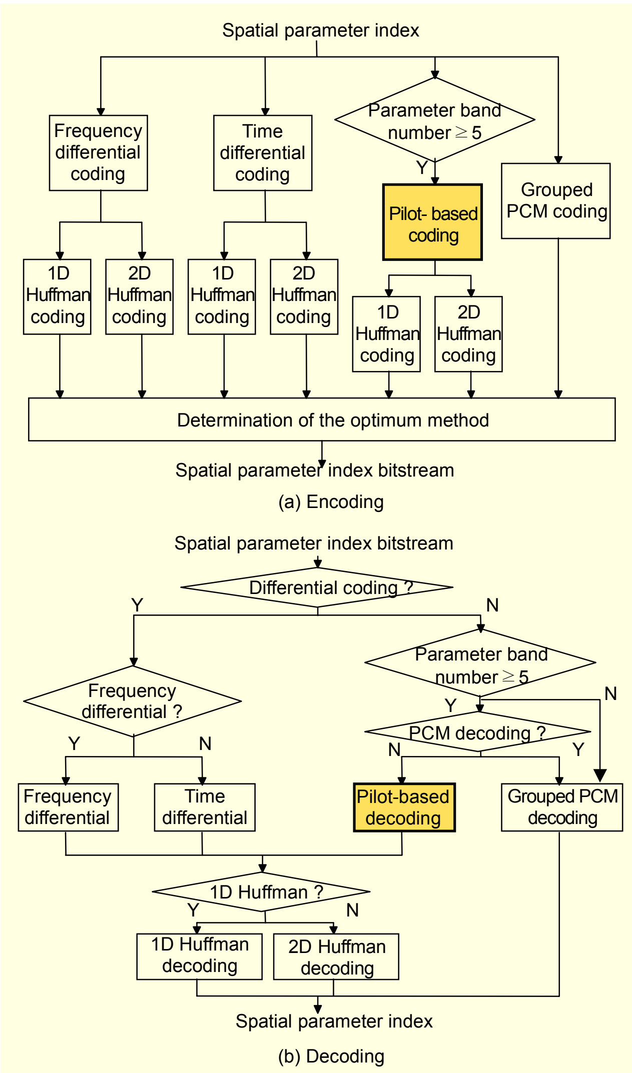


Fig. 1. MPEG Surround encoding and decoding schemes for spatial parameter indexes.

## III. Extended PBC for MPEG Surround

Single PBC, which was adopted for MPEG Surround [7], is advantageous when the variance of data is small. Since the characteristics of audio signals are often dependent on frequency bands, we can make the variance of spatial parameters smaller by dividing the parameter bands into $k$ subgroups and applying PBC to each subgroup. Then, the pilots are determined as

$$(p_1, \cdots, p_k) = \arg\min_{p_m} \sum_{m=1}^{k} \sum_{n=N_m}^{N_{m+1}-1} g(x(n) - p_m), \quad (3)$$

where $p_m$ and $N_m$ are the pilot and the start position of the $m$-th subgroup and $N$ is the parameter band number ($N = N_{k+1} - 1$). Exceptionally, $N$ is the double of the parameter band number for time pair PBC. Since $N + k$ values are encoded for bit stream composition, it especially has advantages when $N$ is large. The most important factor that should be determined is the number of subgroups $k$. In extended PBC, each spatial parameter set has different values of $k$ by pre-determined criteria. The criteria are set statistically to minimize the total bit consumption.

To determine the criteria, we conducted experiments using multi-channel sound databases and calculated the bit consumptions as a function of $k$. Since the bandwidths of the parameter bands are almost logarithmically equivalent [7], we used $N/k$ parameter bands for each subgroup for PBC with multiple pilots. When $k > 1$, the pilots were coded differentially. The results show that the optimum number of pilots is either 1 or 2 depending on the parameter band number (which is one of 4, 5, 7, 10, 14, 20, and 28 [7]) and the spatial parameter type. Using PBC with three or more pilots was disadvantageous due to the bit consumption by the pilots. In Fig. 2, the bit consumption ratios of the dual PBC scheme to the single PBC scheme are depicted. The results were calculated for core spatial parameter data excluding header flags common to both the schemes, which amount to 6 bits per parameter set. Based on the experimental results, we set the rules for extended PBC as follows:

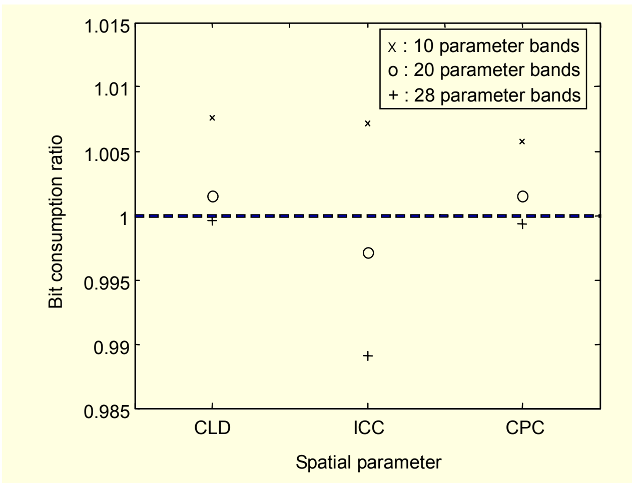


Fig. 2. Bit consumption ratios of the RM0+dual PBC scheme to the RM0+single PBC scheme.

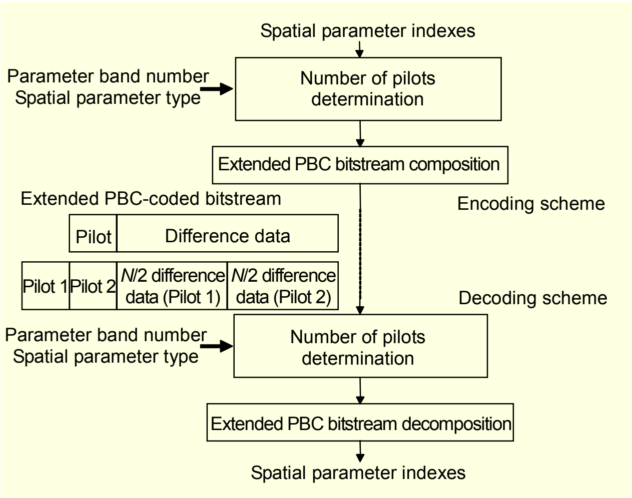


Fig. 3. Extended PBC encoding and decoding schemes for MPEG Surround.

- When the parameter band number is less than 5, PBC is not applied.
- Only single PBC is applicable when the parameter band number is between 5 and 20 for CLD and CPC, and between 5 and 14 for ICC.
- Only dual PBC is applicable when the parameter band number is 28 for CLD and CPC, and either 20 or 28 for ICC.

The overall encoding and decoding schemes for the extended PBC scheme shown in Fig. 3 are similar to those in Fig. 1. In the encoder, the number of pilots is first determined by the rules and then the coding scheme is applied using the same Huffman codebooks used for differential coding. Especially for dual PBC, the second pilot is coded differentially from the first pilot. In the decoder, the number of pilots is first determined by the rules and then the decoding scheme is applied. The spatial parameter indexes are calculated using (2) on a subgroup basis. It should be noted that the extended PBC does not introduce additional loss to the current coding scheme or any additional Huffman tables.

# IV. Experimental Results

Experiments were performed to assess the proposed scheme. We calculated bit consumptions by the original reference model (RM0), RM0+single PBC, and RM0+extended PBC. For comparison, we also added a coding scheme by RM0+dual PBC. We used 11 test items listed in [8], where ACC was used for the encoding and decoding of downmixed sounds. Their lengths were about 20 seconds with sampling frequencies of 44100 Hz for each item. We used 5 modes for the experiments, where TC2 was based on 5-1-5 channel conversions. The other modes were based on 5-2-5 channel conversions

In Fig. 4, the bit reductions by each scheme are depicted, where the results were calculated for the core spatial parameter data. The proposed scheme is superior to the single PBC scheme except for the LR2 mode, where the parameter band number is 10 and the two schemes are identical. The bit reduction by the proposed scheme is more significant for the TC1 and TC2 modes, which use 28 parameter bands, than for the high quality (HQ) and low rate 1 (LR1) modes, which use 20 parameter bands. Note that the bit consumption for the HQ mode is less than those of the TC1 and TC2 modes since the high bit rate of the HQ mode mainly comes from residual coding, which was not included in the results. The proposed scheme has a bit reduction ranging from 1.24% to 2.19% compared to the RM0, and a bit reduction of 0.47% compared to the single PBC scheme for the TC2 mode. The bit reductions are meaningful since the proposed scheme is lossless compared to both the other schemes, which are already highly optimized. It is worth noting that the single and dual PBC schemes are advantageous for low and high bit rate modes, respectively. More detailed results are listed in Table 1.

The bit consumption ratios were calculated for each item for the TC1 and TC2 modes and are depicted in Fig. 5 for more detailed analysis of the results. The extended PBC scheme is superior to the single PBC scheme for most of the items. Exceptionally, both the schemes are comparable for item 1 depending on the mode. The bit reductions come mainly from ICC, although the bit consumption is also reduced for both CLD and CPC. Especially for item 7 in the TC1 mode, the extended PBC scheme has a bit reduction of 2.72% compared to the single PBC scheme for ICC.

The RM0 bit rates, including all the header information, are 11.59 kbps and 11.72 kbps for the TC1 and TC2 modes, respectively. The bit rate reductions for the single and extended

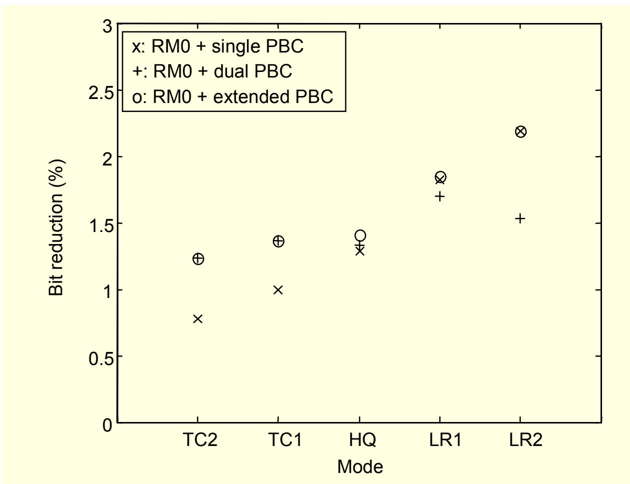


Fig. 4. Average bit reductions of RM0+single PBC, RM0+dual PBC, and RM0+extended PBC compared to RM0.

Table 1. Number of bits used for RM0, RM0+single PBC, RM0+dual PBC, and RM0+extended PBC schemes for 5 modes of MPEG Surround.

| Mode | Para-meter | RM0 | RM0+ single PBC | RM0+ dual PBC | RM0+ extended PBC |
|---|---|---|---|---|---|
| TC2 (5-1-5) | CLD | 1,248,328 | 1,240,978 | 1,240,432 | 1,240,432 |
| | ICC | 856,123 | 847,129 | 838,003 | 838,003 |
| | CPC | . | . | . | . |
| | Total | 2,104,451 | 2,088,107 | 2,078,435 | 2,078,435 |
| TC1 (5-2-5) | CLD | 642,832 | 639,060 | 638,946 | 638,946 |
| | ICC | 578,995 | 572,290 | 565,913 | 565,913 |
| | CPC | 635,500 | 627,507 | 627,121 | 627,121 |
| | Total | 1,857,327 | 1,838,857 | 1,831,980 | 1,831,980 |
| HQ (5-2-5) | CLD | 478,221 | 474,260 | 474,605 | 474,260 |
| | ICC | 435,931 | 429,026 | 427,334 | 427,334 |
| | CPC | 473,908 | 466,974 | 467,709 | 466,974 |
| | Total | 1,388,060 | 1,370,260 | 1,369,648 | 1,368,568 |
| LR1 (5-2-5) | CLD | 347,007 | 338,679 | 339,560 | 338,679 |
| | ICC | 264,825 | 260,556 | 260,297 | 260,297 |
| | CPC | 467,141 | 460,117 | 460,774 | 460,117 |
| | Total | 1,078,973 | 1,059,352 | 1,060,631 | 1,059,093 |
| LR2 (5-2-5) | CLD | 185,415 | 181,140 | 182,513 | 181,140 |
| | ICC | 141,098 | 137,615 | 138,595 | 137,615 |
| | CPC | 239,773 | 235,140 | 236,495 | 235,140 |
| | Total | 566,286 | 553,895 | 557,603 | 553,895 |

PBC schemes are 88 bps and 120 bps for the TC1 mode, and 78 bps and 123 bps for the TC2 mode. Since the combinations of two pilots should be checked for the extended PBC scheme, the complexity is slightly increased compared to the single PBC scheme on the encoder side. However, on the decoder side, the extended PBC scheme additionally requires reading the second pilot once per spatial parameter set compared to the single PBC scheme. For example, the complexity increase is maximally 172 Huffman codebook readings and 172 sign bit readings per second on the decoder side for the TC2 mode shown in Table 1, assuming only one parameter set is present in a frame.

## V. Conclusion

We proposed extended PBC and applied it to MPEG Surround encoding and decoding schemes. Experimental results show that the proposed scheme achieves meaningful bit rate reductions compared to RM0 and RM0+single PBC schemes, where the latter was adopted for MPEG Surround. On the decoder side, the complexity increase of extended PBC is negligible compared to that of single PBC. Since the current Huffman codebooks in MPEG Surround are optimized for differential coding, the bit rate reduction can be further increased if we use PBC-optimized codebooks, which will slightly increase the memory size. The PBC-schemes are applicable to any codecs for lossless bit rate reduction.

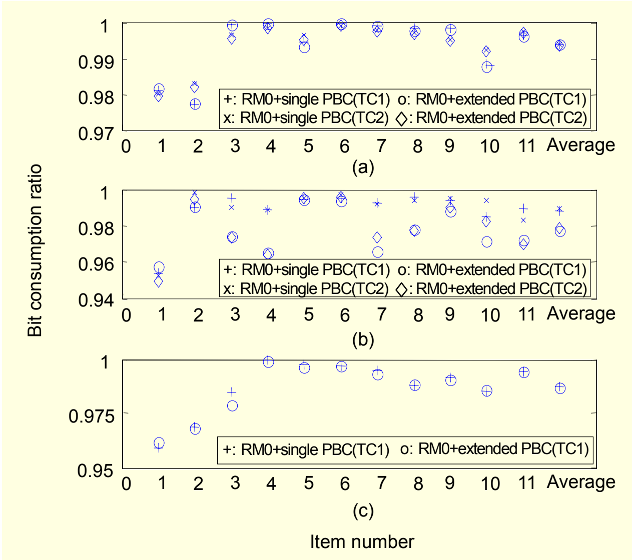


Fig. 5. Bit consumption ratios of the extended PBC and single PBC schemes to RM0: (a) CLD, (b) ICC, and (c) CPC.

## References

[1] M. Bosi et al, "ISO/IEC MPEG-2 Advanced Audio Coding," *J. Audio Eng. Soc.*, vol. 45, no. 10, 1997, pp. 789-814.

[2] M.F. Davis and M.C. Fellers, "Virtual Surround Presentation of Dolby AC-3 and Pro Logic Signals," *AES 103rd Convention*, Preprint no. 4542, 1997.

[3] F. Baumgarte and C. Faller, "Binaural Cue Coding - Part I: Psychoacoustic Fundamentals and Design Principles," *IEEE Trans. Speech Audio Processing*, vol. 11, no. 6, 2003, pp. 509-519.

[4] C. Faller and F. Baumgarte, "Binaural Cue Coding - Part II: Schemes and Applications," *IEEE Trans. Speech Audio Processing*, vol. 11, no. 6, 2003, pp. 520-531.

[5] S. Quackenbush and J. Herre, "MPEG Surround," *IEEE Multimedia*, vol. 12, no. 4, 2005, pp. 18-23.

[6] H.O. Oh, Y.W. Jung, H.S. Pang, D.S. Kim, and J.H. Lim, "Proposed Core Experiment on Pilot-Based Coding of Spatial Parameters for MPEG Surround," *ISO/IEC JTC1/SC29/WG11*, M12549, Oct. 2005.

[7] ISO/IEC JTC1/SC29/WG11, *ISO/IEC FDIS 23003-1:2006, MPEG Surround*, N8324, July 2006.

[8] ISO/IEC JTC1/SC29/WG11, *Call for Proposals on Spatial Audio Coding*, N6455, Mar. 2004.